**The Dissertation Committee for Seung-Eun Chang Certifies that this is the approved version of the following dissertation:**

# THE PHONETICS AND PHONOLOGY OF SOUTH KYUNGSANG KOREAN TONES

**Committee:**

Scott Myers, Supervisor

Megan Crowhurst

Tony Woodbury

Harvey Sussman

Randy Diehl

David Silva

# THE PHONETICS AND PHONOLOGY OF SOUTH KYUNGSANG KOREAN TONES

**by**

**Seung-Eun Chang, B.A., M.A.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**December, 2007**

## Dedication

To my family,

with gratitude and love

## Acknowledgments

My most sincere thanks and gratitude go to Professor Scott Myers, my supervisor and professor. From my first year of graduate school, his insightful comments on my scattered ideas and rough manuscripts have helped me to write cogent and focused analyses and, finally, to complete my dissertation. Additionally, his timely encouragement enabled me get through the difficult times by rekindling my waning energy. Professor Myers, my resolution to do my best came from the desire to make you proud of me. As a linguist and a professor, but most importantly, as a person, you have always been a role model for my academic life. I am truly blessed to be a recipient of your guidance and mentorship.

I also would like to thank five scholars who comprised the rest of my dissertation committee. I am grateful to Professor Megan Crowhurst and Professor Tony Woodbury for their contributions, not just to this dissertation, but also to my training as a linguist through their courses on the phonology and field methods of tone languages. Special thanks to Professor Randy Diehl and Professor David Silva for their useful comments on the statistical analysis and experimental methods. Professor David Silva, your comments written in Korean were practical and interesting. I thank Professor Harvey Sussman for his supporting and helpful remarks on the phonetic aspects of my analysis.

Thanks are also due to Professor Xi Yu, Professor Janet Pierrehumbert, and Professor Robert Ramsey, who enriched my studies by their papers, emails, and comments at various points.

I extend deep thanks to the anonymous subjects who gave of their time to participate in this research and to the ITC research consultants at UT for their prompt and careful consultations on my statistical analysis. Many thanks to Alexis (Hye-Sun) for her warm caring for my baby, when such assistance was needed. I also thank So-Hee for her presence at UT while I was in Austin. I will greatly miss our overnight talks.

My love and thanks go to my parents and family. Their love for me and faith in me were the source of my determination. To my husband, Il-Soo, and my baby, Daniel-Wu, you are the jewels and angels I cherish most. Thanks for being in my life.

I always thank God for standing by me.

# The phonetics and phonology of South Kyungsang Korean Tones

Publication No.______________

Seung-Eun Chang, Ph.D.
The University of Texas at Austin, 2007

Supervisor: Scott Myers

South Kyungsang Korean is spoken in the southeastern part of Korea. The critical problems in analyzing this language are the substantial inconsistency among authors concerning tonal descriptions, and the typologically unusual tone alternation patterns, which have consequently led to vague tonal contrasts and conflicting data. Given that the majority of earlier studies on this language have been based on transcription data, there is limited physical evidence concerning these issues. This research therefore presents the phonetic realization of tonal contrasts, based on F0 production and perception experiments.

It was found that there are three distinct tone classes in monosyllabic roots, and they are different in F0 height, syllable duration, and the timing of F0 peak and fall both in unsuffixed words and in suffixed words. It has been suggested, based on this acoustic evidence, that South Kyungsang Korean has three different tone contrasts in monosyllabic roots, e.g., M, H, and R, and that the different tone alternation patterns of these three tone classes in suffixed words might be the reflection of the phonetic implementation of each tone class.

This work provides not only new basic facts of South Kyungsang Korean tonology, but also another way of understanding tone targets and tone alternation patterns. While clarifying the complications in a particular language, the findings in this work also contribute to debated topics of modern intonational research, concerning such as phonological and linguistic units in the F0 continuum, the temporal alignment of F0 features with segmental strings, and perception of tonal contrasts.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1: Introduction and Background

This study examines the phonetics and phonology of tones in South Kyungsang Korean, spoken in the southeastern part of Korea. The research presents the phonetic realization of tonal contrasts, based on fundamental frequency (F0) production and perception experiments. The investigation was motivated both by specific questions related to the South Kyungsang Korean language and by the theoretical aspects of intonation studies. The critical problem in analyzing this language is the substantial inconsistency among authors concerning tonal descriptions, which have consequently led to vague tonal contrasts and conflicting data in South Kyungsang Korean. Given that the majority of earlier studies on this language have been based on transcription data (Huh 1955, Ramsey 1975, Kim-C 1978, Kim 1996), there is limited physical evidence concerning this issue.

The current study therefore attempts to determine the possible tone classes in the language, by examining tone alternations in suffixed words, as well as in unsuffixed words, through a quantitative analysis of F0 measurements. The tone alternation occurs in different ways, depending on the tone classes, and thus alternation pattern is one property providing crucial information to distinguish the tone classes (Ramsey 1975, Chung 1980, Kim 1996). However, the alternation patterns also have been a stumbling block for formal analysis, requiring typologically unusual tone alternation behaviors (Chung 1980, Kim 1996). New tonal descriptions for the tone classes as well as the tone alternations are provided in this work, and it is argued on this basis that the tone alternation patterns are not random, but systematically reflect the phonetic implementation of each tonal target.

While clarifying the questions in a particular language, this study is also intended to contribute to the topics of modern intonational research under debate, such as phonological and linguistic units in the F0 continuum, the temporal alignment of F0 features with segmental strings, and perception of tonal contrasts. The first part of this chapter provides an overview of South Kyungsang Korean tones, addressing the problems in which the current study is set. The motivation and the organization of the study are presented in the remainder of the chapter.

## 1.1 SOUTH KYUNGSANG KOREAN

South Kyungsang Korean is spoken in the southeastern part of Korea, the lower part of the marked area in Fig 1.1. In this figure, the whole marked area is Kyungsang province, and within that region the upper part (labeled Kyongsang Buk-do) is the North Kyungsang province and the lower part (Kyongsang Nam-do) the South Kyungsang province. Kyungsang Korean has two subdialects, North Kyungsang and South Kyungsang, spoken in their respective areas.

Research on this language has found that there are tonal differences between the two subdialects (Ramsey 1975, Chung 1991, Kim 1996, Kenstowicz & Park 2006), and there is a relatively large body of literature about North Kyungsang Korean (Sohn, 1987, Kim 1988, Chung 1991, Kim-N 1996). Therefore, this dissertation is focused only on South Kyungsang Korean. The relevant differences of two dialects will be discussed in a later part of the chapter. After a brief overview of the phonological inventory of the language, this introduction chapter will turn to a more focused discussion of the tones.


Figure1.1: Map of Korean Peninsula (from http://www.lib.utexas.edu/maps/korea.html)

### 1.1.1 Segments of South Kyungsang Korean

The vowel system of South Kyungsang Korean can be described in terms of six segments as seen in (1) (Ramsey 1975, Chung 1980, Sohn 1999), compared to the overall seven-vowel system of Seoul Korean. Seoul Korean has a high-central vowel [ɨ] as well (Hong 1988).

(1) Vowels of South Kyungsang Korean

| | front | central | back |
|---|---|---|---|
| high | i | | u |
| mid | e | ə | o |
| low | | a | |

The consonant inventory of South Kyungsang Korean consists of 18 segments, and the phonemic inventory is given in (2). The consonant inventory is the same as in Seoul, except that South Kyungsang Korean has [s] instead of the tensed fricative [s'] in Seoul Korean (Chung 1980, Sohn 1999). $C^h$ represents aspirated consonants and C' represents tensed (reinforced) ones.

(2) Consonants of South Kyungsang Korean (Chung 1991, Sohn 1999)

| | labial | alveolar | alveo-palatal | velar | glottal |
|---|---|---|---|---|---|
| stop | p, $p^h$, p' | t, $t^h$, t' | ʧ, $ʧ^h$, ʧ' | k, $k^h$, k' | |
| fricative | | s | | | h |
| nasal | m | m | | ŋ | |
| liquid | | l | | | |

Although South Kyungsang Korean is generally intelligible to speakers of other Korean dialects, it has a number of unique words that are incomprehensible to the speakers of other dialects, such as *məɾak*$^h$*ada* 'scold' (cf. Seoul Korean *k'ujit'a)*, *musi* 'radish' (cf. Seoul Korean *mu*:), and *sik'əphada* 'startled' (cf. Seoul Korean *nollada*) (Sohn 1999, Lee 2000).

### 1.1.2 Tones of South Kyungsang Korean

I now turn to a more careful consideration of tones in South Kyungsang Korean, the focus of this study. South Kyungsang Korean has lexical tones and tone minimal pairs, such as *nuń* (H) 'eye', *nuň* (R) 'snow', while standard Korean does not (Huh 1955, 1985, Ramsey 1975, Lee 1978, Chung 1980, Kim 1996, Cho 1996). Although there is much dispute on how many tone targets exist in this language and how to transcribe the tones of each word, authors agree as to which words belong to the same tone class.

For example, while different authors give different tonal descriptions for different classes of monosyllabic words[1] (Choi 1929, Ramsey 1975, Chung 1980, Huh 1985, Kim 1996), they generally share the idea that there are three distinctions in monosyllabic words, as in (3).

(3) Three classes of monosyllabic words in South Kyungsang Korean

| | | | | | |
|---|---|---|---|---|---|
| a. sul | 'alcohol' | b. mul | 'water' | c. tol | 'stone' |
| mal | 'horse' | nun | 'eye' | pem | 'snake' |
| k'ot | 'flower' | pab | 'steam rice' | pe | 'two times' |
| sot | 'pot' | son | 'hand' | kan | 'liver' |
| son | 'guest' | ot | 'cloth' | mal | 'speech' |

As noted above, the tonal descriptions for these classes are various among authors. I will refer to these three classes of monosyllabic words as M, H, and R, respectively, in this work, because this labeling can reflect, to some extent, my own observations of each tone class. In my pilot study, I listened to items and examined the pitch tracks and spectrograms of extensive controlled samples from several native speakers of South Kyungsang Korean as well as my own speech as a native speaker of this language. The

[1] The various tonal descriptions across authors will be discussed shortly while examining each tone class.

statements including "my observations" or "preliminary observations" henceforth are made based on my perceptions and my examinations of the data in those pilot studies.

Returning to three classes of monosyllabic words, I provide detailed descriptions of each tone class, by reviewing the previous literature and my own observation. The tone alternation patterns in suffixed words are also examined because they have a significant function in distinguishing the tone classes in this language (Ramsey 1975, Chung 1980, Kim 1996).

### *1.1.2.1 M tone class*

The tone class that we call M-class corresponds to the group (a) in (3) above. According to my observation, the items of M-class have a quite high initial pitch and low final pitch when listening to them in isolation.

In suffixed words, when the M-toned root is followed by a monosyllabic suffix or a vowel-initial polysyllabic syllabic suffix, the high pitch occurs in the root, but when the M-toned root is followed by a consonant-initial polysyllabic suffix, the high pitch occurs in the suffix and the root has a low pitch. The high pitch in suffixed words is perceived slightly higher than that in isolation words. I also found that every word has a final low pitch in statement, and thus I assume that there is an L% boundary tone in the language (Cho 1996).

The examples of M-class in suffixed words can be illustrated as in (4). The morpheme boundary is marked by '-' and the suffixes are underlined and italicized. In transcribing the tone of the words, high tone is marked with the (´ ), low tone with (` ), rising tone with (˘ ), and mid tone will be unmarked in this work.

(4) M-class

| | | | | | |
|---|---|---|---|---|---|
| mun | 'door' | sul | 'alcohol' | k$^{h}$oŋ | 'bean' |
| mún-*î* | 'door (nom.)' | súɾ-*î* | 'alcohol (nom.) | k$^{h}$óŋ-*î* | 'bean (nom.) |
| mún-*dò* | 'door also' | súl-*dò* | 'alcohol also' | k$^{h}$óŋ-*dò* | 'bean also' |
| mún-*əɾò* | 'with the door' | súɾ-*esə̂* | 'in alcohol' | k$^{h}$óŋ-*esə̂* | 'in bean' |
| mùn-*bódâ* | 'than door' | sùl-*bódâ* | 'than alcohol' | k$^{h}$òŋ-*bódâ* | 'than bean' |

The sample F0 tracks for the M-class in isolation and in suffixed words, produced by a female native speaker of South Kyungsang Korean, are given in Fig 1.2. The vertical lines in each contour indicate the syllable boundary. The arrows indicate the F0 peak onset (the first point on the highest F0 points) and the peak offset (the first point followed by a string of successively lower values).

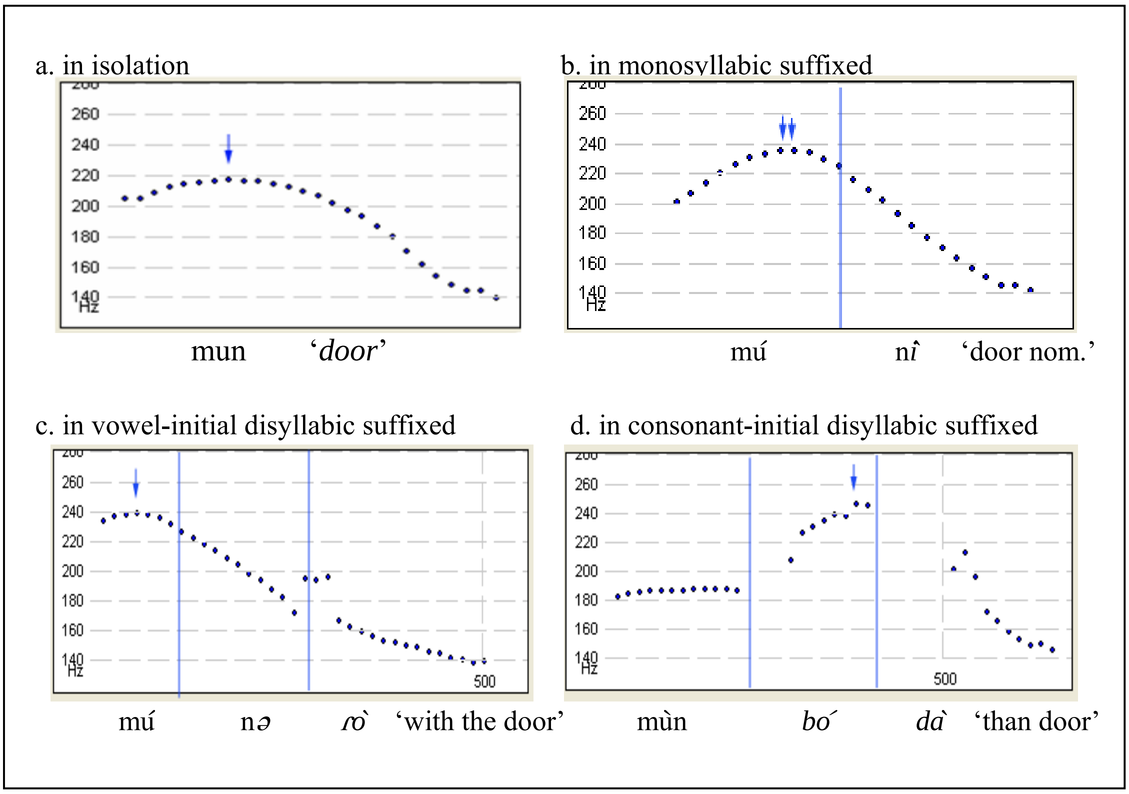



Figure 1.2: F0 contours of *M* –class in isolation and in suffixed words

It can be seen in the figure that the F0 peak comes in the initial syllable in isolation, monosyllabic suffixed words, and vowel-initial polysyllabic suffixed words, and the peak comes in the second syllable in consonant-initial polysyllabic suffixed words. The peak is slightly lower in isolation than in suffixed words, and there are final F0 drops in all words.

The tonal descriptions of this M-class are relatively not conflicting among authors. It has been described as a high tone in most of earlier studies (Choi 1929, Kim-C 1975, Huh 1985, Kim 1996). For example, Choi (1929) provides a detailed description, stating that the initial pitch is high and the final one is low when listening to them in isolation. Also this tone is short, so the final low pitch is barely perceptible (Kim-Y 1986: 3-4, quoting Choi 1929: 97). Although this description is quite similar with my observation and the sample pitch track in Fig. 1.2 (a), the point that this class has a highest pitch among three tone classes of monosyllabic words is conflicting with my observation. This will be discussed further in examining the tone class that we call H-class in next section.

Kim-C (1975) explains tones, comparing them to musical notes, and argues that there are 15 different F0 levels between the highest F0 value and the lowest F0 value. He assigns [1] to the lowest pitch corresponding to the note of C and [15] to the highest pitch corresponding to the note of G. On a scale of 15, high tone can have [12] - [15], and he assigns the high tone to this tone class. His claim is also consistent with Choi (1929) in that this tone class has a highest pitch among three tone classes.

Ramsey (1975) argues that this tone class is atonic. This is based on the tone alternation patterns in suffixed words, that is, although this tone class has a high pitch in isolation, the tone of the root is varied depending on the suffix type. However, the tone alternation pattern that the root has a high tone in suffixed words is consistent with my observation.

Cho (1996), an instrumental study, described this tone class as a "peak-fall F0 contour", by which it could be understood that there is an initial F0 peak and then it is followed by a F0 fall in the pitch contour. This phonetic description is in accord to my observation and the sample F0 track above. Kim (1996) suggests that this tone class is a high tone and the data of tone alternation is consistent with my observation shown in (4). However, Cho (1996) and Kim (1996) are different with my suggestion in that they do not distinguish the M-tone class from H-class.

To summarize, authors generally agree that this tone class has a highest pitch among three tone classes of monosyllabic words. This tone class will be further discussed in comparison with the tone class that we call H-class.

#### *1.1.2.2 H tone class*

The tone class that we call H-class corresponds to the group (b) in (3) above. My observation found that the items of H-class have a quite high initial pitch and low final pitch when listening to them in isolation, and it is almost impossible to distinguish the items of H-class and M-class from listening to them when produced in isolation. Nevertheless, I discriminate them as two different classes mainly based on their different tone alternation patterns in suffixed words, and partly based on their small, but consistent differences in F0 values and syllable durations in F0 tracks.

In suffixed words, when the H-toned root is followed by a suffix, the initial F0 in the first syllable is rather high but the peak occurs in the first syllable of the suffix, regardless of suffix type. This can be transcribed as in (5), and the sample F0 tracks for H-class in isolation and in suffixed words are given in Fig. 1.3.

(5) H-class

| | | | | | |
|---|---|---|---|---|---|
| nún | ‘eye’ | múl | ‘water’ | són | ‘hand’ |
| nun-*í* | ‘eyes (nom.) | muɾ-*í* | ‘water (nom.) | son-*í* | ‘hand (nom.) |
| nun-*dó* | ‘eyes also’ | mul-*dó* | ‘water also’ | son-*dó* | ‘hand also’ |
| nun-*ɔ́ɾɔ̀* | ‘with eyes’ | muɾ-*ésə̀* | ‘in water’ | son-*ɔ́ɾɔ̀* | ‘with a hand’ |
| nun-*bódà* | ‘than eyes’ | mul-*bódà* | ‘than water’ | son-*bódà* | ‘than hand’ |

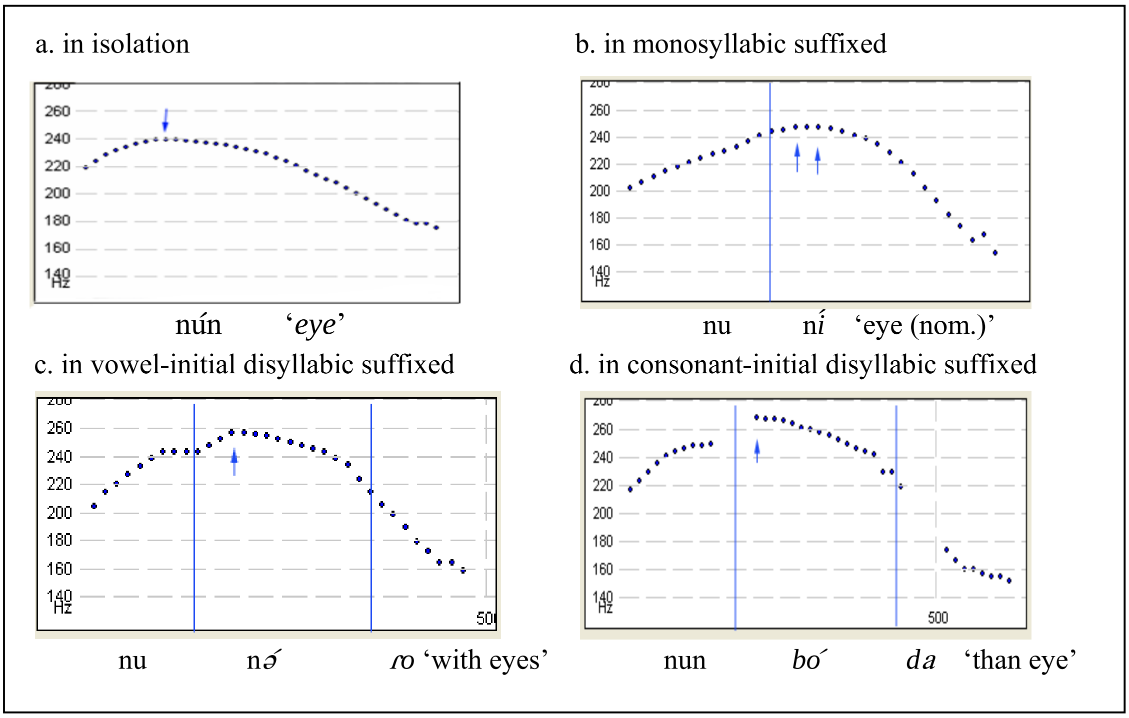


Figure 1.3: F0 contours of *H* in isolation and in suffixed words

It is seen in the figure that the F0 contour in isolation is quite similar with that of M-class, but the peak F0 values are vaguely higher than that of M-class. Therefore, this tone class is referred to as H-class in this work. In addition, the syllable duration of this tone class is shorter than for M-class. It is also shown in suffixed words that the peak occurs in the onset of the second syllable with F0 rising through the initial syllable. This could easily be interpreted as tone spread, since both the root (initial syllable) and its successor include regions of high F0 values. However, although the initial F0 values in the first syllable are relatively high, the F0 peak usually occurs only in the second syllable in my preliminary observation.

The tonal descriptions of this tone class are inconsistent in some way among authors. Choi (1929) explains that it is a mid tone, which is neither high nor low, neither long nor short, and Ramsey (1975) explains that this tone type is tonic, which has a high pitch irrespective of suffix type.

Some researches providing the phonological analysis (Chung 1980, Kim 1996) and phonetic data (Cho 1996) do not separate this tone class from M-class, as noted before. In other words, Chung (1980) and Kim (1996) consider both M-class and H-class are high tones, and they are distinguished only by their tone alternation pattern in suffixed words. Similarly, Cho (1996) classifies the words of M-class and H-class into same category "peak-fall F0 contour".

As for the tone alternation patterns in suffixed words, it was unanimously suggested in earlier studies that a high tone is associated with both the first and the second syllable in suffixed words (Ramsey 1975, Kim 1996). According to these descriptions, H-class in monosyllabic suffixed words is represented as HH (e.g., nɨ́n+ɨ́ '*eye (nom.)*). This perception might be possible because the F0 values are high both in the initial and the second syllables. However, the peak regularly occurs in the second syllable rather than it is stretched over two syllables with a long peak plateau, as seen in Fig. 1.3.

It is interesting that the class we call H-class has been described to be lower than the class we call M-class, in some studies (Choi 1929, Kim 1980). In fact, my observations found that M-class and H-class in isolation are hardly distinguishable from listening to them. Perhaps, therefore, that the distinction between them was to a great extent determined by the pitch and tone alternation patterns in suffixed words by authors. Hence, the answer for the question why the class we call H was perceived lower than the class we call M by some authors can also be found in suffixed words.

My conjecture is that it might be due to the slightly lower F0 value of the initial syllable, i.e., the root, for H-class in suffixed words. Since the F0 peak is realized in the second syllable, i.e., the suffix, for H-class, the F0 values in the root is relatively low. On the other hand, the peak is in the root for M-class, and thus the F0 values in the root are quite high for M-class. Therefore, it is possible that the F0 height is perceived to be higher for M than for H, when considering the F0 values of the root in suffixed words. Some authors might want to distinguish these two tone types, depending on the pitch in suffixed words, because the two tones are almost the same to one's ears in isolated words. This might have led them to argue that the class we call M is a high tone and the class we call H is a mid tone.

#### *1.1.2.3 R tone class*

The tone class that we call R-class corresponds to the group (c) in (3). My preliminary observation found the tone class we call R-class has a dramatically lower initial pitch, a slightly lower and later peak pitch than the other classes do. In addition, this tone class is longer than the others, probably because both the initial low pitch target and the peak pitch target need to be realized in a syllable.

The alternation in R is distinct from either M or H, i.e., the peak occurs in the final syllable of the suffix, and the root has a low pitch. The peak is perceived to be higher in suffixed words than in isolation. This tone alternation can be illustrated as in (6), and the sample F0 tracks for R in isolation and in suffixed words are given in Fig. 1.4.

(6) R-class

| | | | | | |
|---|---|---|---|---|---|
| nǔn | ‘snow’ | kǎn | ‘liver’ | tǒl | ‘stone’ |
| nùn-*í* | ‘snow (nom.) | kàn-*í* | ‘liver (nom.) | tòɾ-*í* | ‘stone (nom.) |
| nùn-*dó* | ‘snow also’ | kàn-*dó* | ‘liver also’ | tòl-*dó* | ‘stone also’ |
| nùn-*əɾó* | ‘with snow’ | kàn-*əɾó* | ‘with liver’ | tòɾ-*esə́* | ‘from stone’ |
| nùn-*bodá* | ‘than snow’ | kàn-*bodá* | ‘than liver’ | tòl-*bodá* | ‘than stone’ |

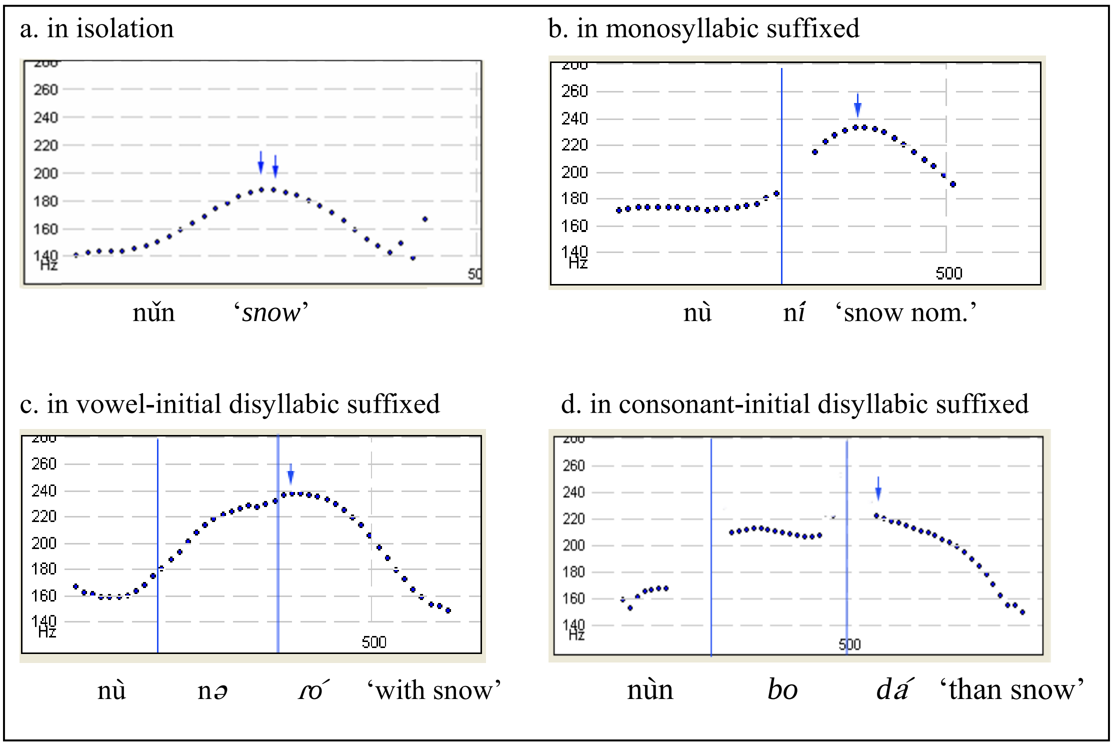



Figure 1.4: F0 contours of *R* in suffixed words

It is seen in isolation words that R-class has a F0 peak that is slightly later and lower than in H-class, and the initial F0 value is also lower than the other two classes. In addition, the syllable duration is longer compared to other classes.

In suffixed words, the F0 rise comes in the initial syllable, and the peak and fall come in the second syllable in monosyllabic suffixed words, as in Fig 1.4 (b). On the other hand, the F0 rise comes in the initial syllable, and the peak and fall come in the third syllable in disyllabic suffixed words, as in (c,d).

The tonal descriptions of this tone class are the most conflicting among authors. Choi (1929) provides careful description that the initial pitch of this tone is low, but the pitch rises above mid-tone in the later portion, and it is long. Although Choi (1929) refers to this tone class as a low tone, his description is similar with my observation. Ramsey (1975) explains that it is a rising tone. Cho (1996) distinguishes this tone type from the earlier two tone types, by describing it as a “rise-fall F0 contour”, by which it could be understood that there is an F0 rising and then it is followed by a F0 fall in the pitch contour. These all descriptions are quite consistent with my observation and the pitch track in Fig. 1.4 (a).

On the other hand, it has been transcribed as a level low tone in Kim (1980) and Huh (1985). In contrast, Kim (1996) assumes that this is also a high-toned long vowel that is same with the previous two tone classes, and that they are distinct in syllable duration and alternation form in suffixed words. Also, his tone alternation pattern in disyllabic suffixed words is different with my suggestion, that is, the two syllables of disyllabic suffix have been represented to be associated with a high tone (e.g., nu+nə́+ɾó ‘with snow’). This perception might be possible because both the second syllable and third syllable include regions of high F0, as seen in Fig. 1.4 (c).

Having reviewed the monosyllabic words, we now turn to the disyllabic words that also have a different tone alternation patterns in suffixed words.

#### *1.1.2.4 Disyllabic words*

Authors agree that there are three or four tone classes in monomorphemic disyllabic words as in (7), although their tonal transcriptions are different among authors (Huh 1955, Ramsey 1975, Cho 1996, Kim 1996).

(7) Disyllabic words in South Kyungsang Korean

| | | | |
|---|---|---|---|
| a. ane | 'wife' | b. nalge | 'wing' |
| moɾe | 'sand' | moɾe | 'the day after tomorrow' |
| nabi | 'butterfly' | kuɾəm | 'cloud' |
| kəul | 'mirror' | mogi | 'mosquito' |
| koɾi | 'ring' | $p^h$aɾi | 'fly' |
| c. namu | 'tree' | d. saɾam | 'people' |
| taɾi | 'bridge' | pabo | 'fool' |
| paɾam | 'wind' | imɟa | 'owner' |
| əlgul | 'face' | ange | 'fog' |
| namul | 'wild vegetable' | tambe | 'cigarette' |

I will refer to these tone classes (a, b, c, and d) as HL, HH, LM, and LH, respectively, based on my preliminary observation. It was observed that the peak pitch is in the first syllable, and low one is in the second syllable for the items in (a), and thus I refer to this tone class as HL. The peak pitch is in both syllables for the items in (b) and thus I refer to this tone class as HH. As for the items in (c) and (d), the initial pitch is very low and the second syllable has a peak. Although I perceived the items in (c) and (d) as almost the same when the word is produced in isolation, I discriminate them as LM and LH, respectively, based on their different tone alternation patterns in suffixed words and small different peak F0 values in pitch track.

The sample F0 tracks for the four tone types of disyllabic words are given in Fig. 1.5. Fig. 1.5 (a) is for HL, (b) HH, (c) LM, and (d) LH. F0 peak realizing a high tone occurs in the first syllable, and the F0 fall comes also in the same first syllable for the tone class that we call HL. On the other hand, the peak begins in the first syllable and is extended to the second syllable with a long peak plateau for the tone class we call HH. Regarding the tone classes we call LM and LH, the F0 rise begins at quite low value in the first syllable, and F0 peak and fall comes in the second syllable. Although the peak is a little higher for LH than for LM, the difference is quite small.

It is also seen that the F0 peak in HL is exclusively higher than that in others. This pattern can be interpreted as a phonetic effect of anticipatory raising or H-raising, which refers to the phenomenon that the F0 height of a tone is raised when followed by a tone that has a low pitch value. This tendency has been reported for some other languages such as Thai, Mandarin, and Yoruba (Gandour et al., 1994, Xu 1997, Laniran & Clements 2003).

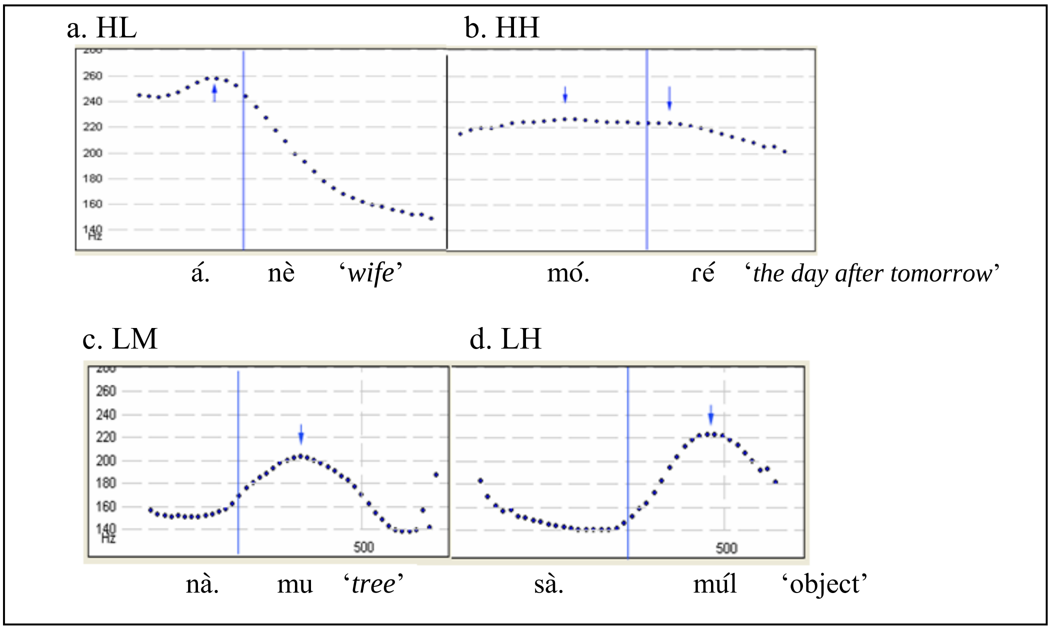



Figure 1.5: Sample F0 contours of disyllabic words

Probably due to this relatively lower peak F0 values for the tone class we call HH than the tone class we call HL, some authors transcribe the tone class we call HH as MM (Huh 1985, Kim-Y 1986). As stated above, however, I interpret this pattern that the peak is higher for the tone class we call HL than the other tone class rather than that the peak is lower for the tone class we call HH than HL, because the peak F0 value in HL is exclusively higher than the peak in other tone classes. As for the tone classes we call LM and LH, most authors do not distinguish LM from LH, as their pitch is heard to be quite similar (Cho 1996, Kim 1996). However, this blurred distinction between LM and LH can be obvious in suffixed words.

When LM is followed by a monosyllabic suffix or a vowel-initial polysyllabic syllabic suffix, the peak occurs in the root, but when LM is followed by a consonant-initial polysyllabic suffix, the peak occurs in the suffix, as in (8a). This alternation pattern is identical with the tone class we call M-class. In contrast, when LH is followed by a suffix, the peak occurs in the first syllable of the suffix, regardless of suffix type, as represented in (8b). This pattern is already seen in the tone class we call H-class.

(8) a. LM

| | | | |
|---|---|---|---|
| nàmu | ‘tree’ | kàmɟa | ‘potato’ |
| nàmú-*nən* | ‘tree (topic.) | kàmɟá-*nən* | ‘potato (topic.) |
| nàmú-*esə* | ‘from the tree’ | kàmɟá-*esə* | ‘from the potato’ |
| nàmu-*bóda* | ‘than tree’ | kàmɟa-*bóda* | ‘than potato’ |

b. LH

| | | | |
|---|---|---|---|
| sàmúl | ‘object, thing’ | nèmó | ‘square’ |
| sàmuɾ-*ə́n* | ‘object (topic.) | nèmo-*nə́n* | ‘square (topic.) |
| sàmuɾ-*ésə* | ‘from the object’ | nèmo-*ésə* | ‘from the square’ |
| sàmul-*bóda* | ‘than the object’ | nèmo-*bóda* | ‘than the square’ |

The sample pitch tracks of LM and LH with a monosyllabic suffix are given in Fig. 1.6. It is seen that the F0 peak is realized in the middle portion of the second syllable in LM, but it is at the onset of the third syllable in LH. The division between LM and LH based on this alternation patterns has not been addressed in earlier studies, and it is first reported in this research.

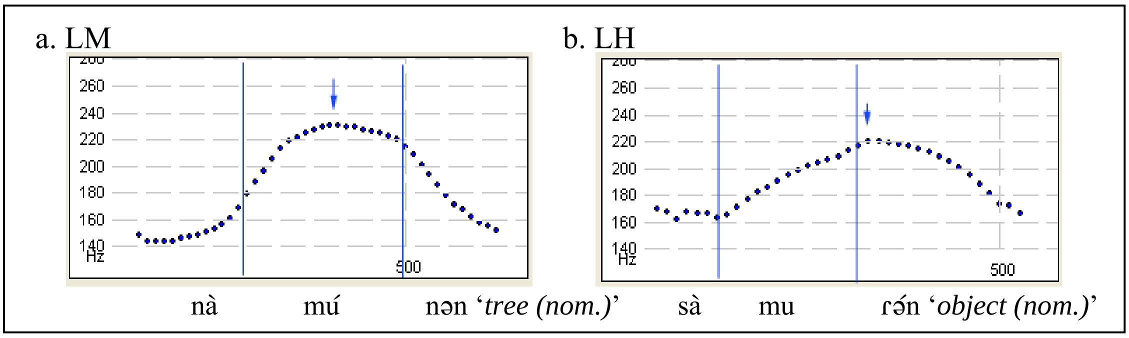



Figure 1.6: F0 contours of LM and LH in monosyllabic suffixed words

## 1.2 Motivation of the Study

### 1.2.1 Problems with the previous studies

The discussion so far has illustrated that there is a considerable gap between my preliminary observation and previous studies, regarding both the tone contrasts in unsuffixed words and tone alternations in suffixed words. While the earlier studies have provided extensive and well considered explanations for tonal contrasts, their transcriptions vary from author to author, even for the same material. This problem typically happens in tonal research based on transcriptions. Myers (2003) has pointed out that, in Kinyarwanda, for example, there are five different tone transcriptions for the word *umusore* “youth” in five different publications. He states that this word has been described as *umusóre* in Sibomana (1974), *umǔsôre* in Overdulve (1975), *umúsôre* in

Kimenyi (1979), *umúśóre* in Faurere and Rialland (1985), and *umūsóre* in Jouannet (1989) (Myers 2003:73). The possible reason is that their descriptions have been influenced by the authors' subjective judgment as well as their theoretical assumptions. Labov (1994) discusses the problem of subjective biases in transcription data, positing that, because the most skilled transcribers of impressionistic data are already informed by the theoretical issues they are studying and influenced by the phonemic categories of their own dialects, it is inevitable that such knowledge will affect their subjective judgments.

Another problem is that, due to their subjectivity, transcription data are not replicable. If there is inconsistency among transcriptions, and one has doubts about generalizations based on those data, one cannot test the generalization by redoing the transcription. However, if one defines a series of experiments with enough precision, then any other observer following the procedure would end up with the same measurement.

A well-defined measure does not vary significantly from observer to observer. Therefore, experimental methods have a great advantage over transcriptions in accuracy, reliability and, most importantly, objectivity. Further, it has been shown that in the sample pitch tracks that the differences between monosyllabic words take up three dimensions (F0 peak location, F0 height and duration), but the differences are quite small. This would explain why these tone contrasts have led to various transcriptions in this language. Therefore, another advantage of the experimental approach is that it provides data as gradient and as fine-grained as the tone target and tonal contrasts.

Cho (1996) is a phonetic study that provides the pitch descriptions of tone types in South Kyungsang Korean. He discusses the different tone types by employing the phonetic interpolation between a lexically linked high tone (H) and a boundary low tone (L%). Since it was also found in my observation that every word has a final low pitch in statement, it was already noted that we assume there is his boundary low tone (L%) in this language. However, while Cho (1996) offers the first acoustic data for this language, his study does not examine the tone categorization in a thorough way because his study mostly deals with the phrasal intonation rather than the tone categorization of the word itself. For example, he assumes that the tone class we call M-class and H-class are same tones and classifies them into one tone class, so-called "peak-fall contour", but he does not even present the sample pitch track for each tone class of monosyllabic words, probably because it is not the main concern of his study. He provides one sample pitch track for each word of disyllabic, trisyllabic, quadrisyllabic words, produced by one native speaker of South Kyungsang Korean, but these were not either based on an experiment with multiple subjects and quantitative analysis.

Chang (2003, 2005) also provides instrumental data for tone types in North Kyungsang Korean, a subdialect of Kyungsang Korean. The literature shows that the inventory of tone classes in North Kyungsang Korean is different from that of South Kyungsang Korean. Ramsey (1975: 87) notes quoting Mun (1962), who states that the accents of South and North Kyungsang province are different in the "rising tone" reflex, and he discusses the rising tone only in the context of South Kyungsang Korean. Kenstowicz and Park (2006) mention that, in monosyllables, the South Kyungsang Korean preserves a high vs. rising contrast that corresponds to a short vs. long in the North Kyungsang dialect, suggesting that the rising tone exists only in South Kyungsang Korean. In addition to monosyllabic words, authors (Kim-Y 1986, Chung 1991, Kim 1996, Kim-C 2003) have pointed out the differences in the tonal inventory of two dialects, for example, the "final double high pattern" is possible only in South Kyungsang Korean.

Nevertheless, Chang's (2003, 2005) body of work, which underlies the current study, provides the basis of many of the decisions and questions in this study. Thus, a brief summary of these studies and the issues left unaddressed is provided below.

Chang (2003) was motivated to provide a description of the long vowels in North Kyungsang Korean by the debate in which some researchers argue that every long vowel is an H tone associated with two moras (Kim 1988, Kim 1997, Kenstowicz & Sohn 1998), while others argue that the long vowel in the word initial is a rising tone and, in the word final is a falling tone (Chung 1974, Chung 1991, 2002). The results of this study showed that short and long vowels are different both in length and F0 levels in all four speakers. Low and peak F0s were significantly lower in a long vowel than in a short vowel, and syllable duration was longer in a long vowel than in a short vowel. Although it was concluded that the length variations have different F0 contours as well as length, the tonal nature of the long vowel was not determined in the study.

Chang (2005), therefore, investigated the F0 timing effects of tone classes, expanding the tone classes and phrase positions with six native speakers in the experiment. While that work generally found that the study of the F0 timing can provide crucial evidence for determining the tone classes, the results raised many questions regarding tonal contrasts which the current study will explore and answer.

For example, there was some difficulty in arguing that a high tone is associated with both syllables for the tone class that has been transcribed in HH in transcription based studies (Chung 1991, Kim 1997), because a long peak plateau laid upon two syllables was not consistently found. And the question of why the peak came later in tone class transcribed as HH than in HL was raised, but not resolved in that study. As such, the pilot studies (Chang 2003, 2005) not only failed to distinguish all tone classes in Kyungsang Korean, but also recognized the significant gap between the tone labeling based on transcription data and its phonetic realization in F0 contours in Kyungsang Korean.

### 1.2.2 Goals of the Study

The current study therefore attempts to fill that gap in the literature and determine the tone classes by examining tone alternations in suffixed words, as well as in isolated words. The tone alternation occurs in different ways, depending on the tone classes as seen earlier, and thus tone alteration is one property providing crucial information to distinguish the tone classes (Kim 1988, Chung 1991, Kim 1996, Kim 1997, Kenstowicz & Sohn 1998). The research consists of three production study and two perception experiments.

The broad goal of the production experiment is to determine whether there is any reliable difference in physical realization between tone contrasts, and what these differences are. The examination is focused on three tone classes of monosyllabic words (so-called M, H, and R), given that tone descriptions of monosyllabic words in isolation are the most controversial among authors, and their descriptions in suffixed words are not supported by phonetic data in my pilot studies and preliminary observations. The monosyllabic words, in fact, mean monosyllabic roots because the data will include suffixed words that are disyllabic and trisyllabic.

It was already seen that those three tone classes are different in initial F0, peak F0, the timing of F0 peak and fall, and syllable duration. Therefore, the detailed measurements will include those acoustic factors.

Among three tone classes, the tonal description of the tone class we call R is especially controversial and the description of the tone class we call M is quite consistent among authors. Therefore, the first production experiment includes the comparison between M and R to see how the R-class is different than the M-class in terms of the acoustic factors noted above.

In addition, it was stated that the difference between M and H was in particular vague and their distinction is not consistent among authors. Hence, the second production experiment investigates the comparison between M and H to see whether and how two classes are different each other in aspects of F0 measurements and syllable duration.

The monomorphemic disyllabic words are also investigated because monosyllabic roots in suffixed words are usually transcribed in a same way with the monomorphemic disyllabic words. For example, when the tone class we call H is followed by a monosyllabic suffix, it is transcribed in H+H, which is same with monomorphemic disyllabic words (HH) (Kim 1996). However, the phonetic data in preliminary observation found there are differences in F0 contours between those morpheme types. Therefore, the third production experiment explores the comparison between HL, HH, M+suffix and H+suffix.

In order to test whether listeners of South Kyungsang Korean use those acoustically determined differences in making word identification decisions, two perception tests, i.e., (i) M and R contrast, and (ii) HL and HH contrast, are conducted. The experiment tests the perception of isolated synthetic stimuli in which significant acoustic cues have been systematically manipulated relying on the production studies. The summarized outline of the study is provided in the next section.

### 1.3 OUTLINE OF THE STUDY

The chapters that follow present five phonetic experiments (three production and two perception tasks) designed to compare a series of tone classes. Chapter 2 compares the tone class we call M-class and R-class. Chapter 2a presents the production study of a contrast between M and R in three suffix types (unsuffixed, monosyllabic suffixed, disyllabic suffixed). The differences are tested in terms of F0 peak delay, fall delay, peak plateau, low F0, peak F0, and syllable duration. Chapter 2b provides a perception experiment which tests the claims made in the preceding production experiment. It tests the perception of isolated synthetic stimuli, in which the crucial acoustic parameters (i.e., F0 peak delay, initial F0, and syllable duration) have been systematically manipulated. It has been generally shown that those acoustic cues contribute to the perception of M and R contrast, with a relatively great between-subject and within-subject variations.

Chapter 3 investigates the relationship between tone classes we call M-class and H-class. Chapter 3a provides the production experiment of M and H contrast in three suffix types. The differences are tested in the light of F0 peak delay, fall delay, peak plateau, peak F0 value, and syllable duration. Chapter 3b examines M and H with a monosyllabic suffix (M+suffix, H+ suffix) in comparison with monomorphemic HL and HH. Chapter 3c provides the perception experiment which tests the claims made in the preceding production experiment. It tests the perception of isolated synthetic stimuli, in which the peak plateau, the most critical factor, has been manipulated. This confirms the finding of the production study that the peak plateau difference alone can discriminate between the monomorphemic HL and HH in South Kyungsang Korean.

Chapter 4 summarizes the main findings and arguments of each chapter and the contributions of the study. Moreover, the chapter provides some directions for a follow-up study.

# Chapter 2: Experiment 1

# M and R Contrast

Experiment 1 is concerned with the tone class we call M and R contrast and consists of one production study (Experiment 1a) and one perception study (Experiment 1b). Experiment 1a provides the acoustic data for 6 native speakers of South Kyungsang Korean, comparing the F0 timing and scaling, and syllable duration between M and R. Based on the findings of production study, Experiment 1b tests the perception of isolated synthetic stimuli in which acoustic cues (F0 timing, scaling, and syllable duration) were manipulated. 10 native speakers of South Kyungsang Korean are asked to identify the stimuli between M and R.

## 2.1 Experiment 1a: M and R contrast in Production

The main goal of Experiment 1a is to determine whether there is any reliable difference between two tone classes (we call M and R) in physical realization, and what that difference is.

The transcription of the tone class we call R has been described as a low tone (Huh 1985, Kim 1986), or as a high tone (Chung 1980, Kim 1996), or as a low-mid rising tone (Kim 1986), or as a low-high rising (Chung 1986). On the other hand, most authors suggest that the tone class we call M is a high tone, which has the highest pitch in this language. Note that some authors, such as Kim (1996), argues that both our M and R are same high tone.

However, authors all agree that these two tone types have different alternation patterns in suffixed words. My observation found that when a root of the M-class is followed by a monosyllabic suffix or a vowel-initial polysyllabic suffix, the F0 peak occurs in the root, but when M-toned root is followed by a consonant-initial polysyllabic suffix, the F0 peak occurs in the suffix and the root has a low tone, as in (9a). In comparison, when R is followed by a suffix, the high tone occurs in the final syllable of the suffix, and the root has a low' tone, regardless of suffix type, as in (9b). The morpheme boundary is marked by '-', and the suffix is underlined.

| (9) a. M | | b. R | |
|---|---|---|---|
| mun | 'door' | nǔn | 'snow' |
| mún-ɨ̀ | 'door (nom.)' | nùn-ɨ́ | 'snow (nom.) |
| mún-əɾò | 'with the door' | nùn-esə́ | 'in snow' |
| mùn-bódà | 'than door' | nùn-bodá | 'than snow' |

However, the pattern shown in (12) is also partially incompatible with the previous transcriptions (Chung 1980, Kim 1986, Kim 1996). For example, Kim (1996) argues that the high tone is associated with two syllables of the suffix when R is followed by a disyllabic suffix. Thus R with a disyllabic suffix is described as *nùnésə́* 'in snow' and *nùnbódá* 'than snow' in Kim 1996.

The current experiment attempts to clarify this dispute through a quantitative analysis of F0 measurements. We've already seen in the introduction chapter that the preliminary observations found that R has a lower initial and peak pitch, and longer syllable duration than M in unsuffixed words. Therefore, we can hypothesize that the low F0 and peak F0 values would be lower for R than for M. In addition to my observation, R has been also referred to as a long vowel in earlier studies of this language (Choi 1929, Cho 1996, Kim 1996). Hence, it is expected that the syllable duration would be longer for R than for M. Further, my observation found that R has a slightly later peak than M does.

R might have two tone targets (L+H) in a syllable and the initial low pitch target has to be realized before the second high pitch target. Hence, it is expected that peak would be later for R than for M.

The difference in suffixed words can also be tested with special reference to the timing of peak and fall. If the high tone is associated with a particular syllable, it is expected that the F0 peak will be realized with respect to that syllable (Bruce 1977, Pierrehumbert 1980, Steele 1986, Silverman & Pierrehumbert 1990, Prieto et al. 1995, Arvaniti et al. 1998, Xu 1999, Myers 1999, 2003). If the high tone is in the first syllable for M and with the second syllable for R in monosyllabic suffixed words, for example, then the F0 peak would be realized later for R than for M. If the high tone is associated with two syllables, the peak would be realized as a long peak plateau, stretched onto two syllables, and thus the timing of the peak would be different with the timing of the fall. Hence, if the high tone is associated with two syllables for R with a disyllabic suffix, as suggested in Kim 1996, then the peak plateau would be longer for R than for M, and the F0 fall also would be later for R than for M. However, if the high tone is in one syllable, as proposed in this study, then the peak and fall would come in the same syllable, and thus the timing of the peak and fall would be quite similar and the peak plateau would not be long.

To sum up, it will be tested on this basis whether the class we call M and R are different in low F0, peak F0, syllable duration, and peak and fall delay. The proposed hypotheses for this experiment can be summarized, as in (10), and the proposed difference in the peak location can be schematized, as in Fig. 2.1. The vertical lines in the figure mark the syllable boundary.

(10) Hypotheses

A. The timing of the F0

a. the *F0 peak* and *fall* are attained later in the word in the R-class than in the M-class across suffix type

b. the *F0 peak* and *fall* are attained later in the disyllabic suffixed words than in the monosyllabic suffixed words, and later in the monosyllabic suffixed words than in the unsuffixed words for R class

B. F0 values

c. the *low F0* (F0 value at the lowest point) and *peak F0* (F0 value at the highest point) is lower in the word in the R-class than in the M-class

C. Syllable duration

d. the *test syllable (the first syllable) duration* is longer in the word in R-class than in the M-class

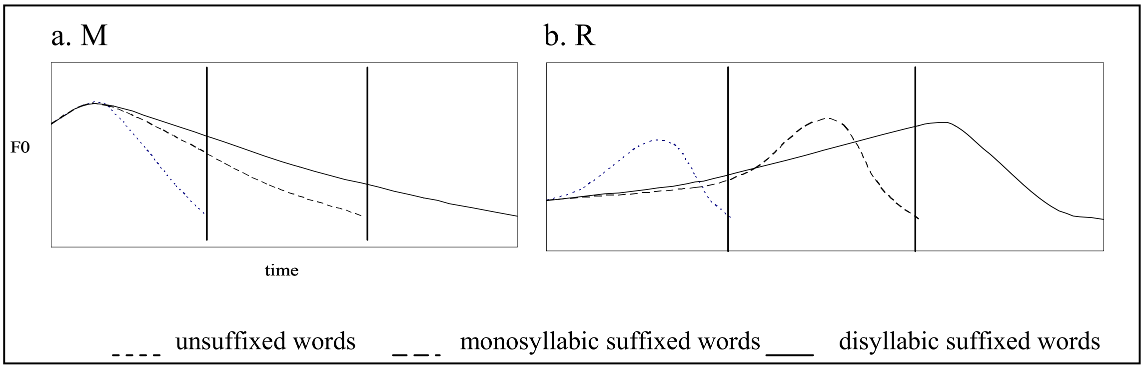



Figure 2.1: The schematic illustrations of F0 contours for M and R in three suffix types

### 2.1.1 Methods

#### *2.1.1.1 Materials*

In order to test these hypotheses, the following factors are included in the experiment: (i) tone type: M and R, (ii) suffix type: unsuffixed, monosyllabic suffixed, and (vowel-initial) disyllabic suffixed. There are six conditions (2 tone type * 3 suffix type) and 10 repetitions for each condition, yielding a total of 60 tokens per speaker.

It is well established that certain factors affect the F0 contours and timing, and thus these factors were manipulated. Sonorants are known to present the least disturbance and interruption of the continuity of F0 contours (Lehiste 1970, Hombert 1974) and vowel-nasal boundaries can be seen clearly both in the waveform and in the spectrogram. Therefore, all test syllables consisted of a nasal followed by a low vowel, and every test syllable was followed by a syllable beginning with a tap in this study.

Phrase position also affects F0 timing. F0 peak has been shown to come earlier in a phrase-final position than in other phrase positions (Steele 1986, Silverman & Pierrehumbert 1990, Myers 2003, Chang 2005). Also, when produced in context, the tonal contours undergo certain variations depending on the adjacent tones (Hyman & Schuh 1974, Gandour et al. 1992, Xu 1997). In order to avoid the phrase-final effect and reduce the influence of adjacent tones, test syllables were always followed by more words in the same phrase. For example, the tone of the following words in each sentence was LH or LHL.

In addition, the quantitative measure of the F0 timing will be the peak delay divided by the "test syllable duration", and we have a hypothesis about syllable duration, so the intrinsic vowel duration effects, according to vowel height (Peterson et al. 1960; Lindblom 1968) and syllable structure (Delattre 1966, Maddieson 1985), need to be controlled. Therefore, the vowel in test syllable was controlled to [a] in the first syllable. Additionally, simple CV syllables were used for every test syllable. The isolated word (e.g., mál) is CVC, but the coda is resyllabified to the next syllable when followed by a vowel in Korean; therefore, every test syllable is CV structured.

Due to the difficulty in finding words that meet all these requirements, one word per each condition was chosen. The sentences used as recording material are presented in Table 2.1. The test word is marked in bold, word boundary is indicated with the space, and the morpheme boundary is with the '-'.

| M | R |
| --- | --- |
| a. unsuffixed | a. unsuffixed |
| **mal** àn-dá<br>*horse know-dec.*<br>'*I know a horse.*' | **mǎl** àn-dá<br>*speech know-dec.*<br>'*I can speak.*' |
| b. monosyllabic suffixed | b. monosyllabic suffixed |
| **máɾ-ì** bò-ín-dà<br>*horse-nom. see-pass- dec.*<br>*'A horse is seen.'* | **màɾ-í** màt-k'é-t'à<br>*speech-nom. right- would be-dec.*<br>*'The speech would be right.'* |
| c. disyllabic suffixed | c. disyllabic suffixed |
| **máɾ-imə̀n** dè-gé-t'à<br>*horse-if enough-would be- dec.*<br>*'A horse would be enough.'* | **màɾ-imə́n** dè-gét-t'à<br>*speech-if enough-would be-dec.*<br>*'The speech would be enough.'* |

Table 2.1: Test materials[2]

### *2.1.1.2 Speakers*

Six adult native speakers (aged 24~58, three females and three males) of South Kyungsang Korean produced the materials. All speakers are linguistically naïve, and they were born and grew up in Pusan, the capital of South Kyungsang. Four of them lived only

2 *Pass* is an abbreviation for passive morpheme, *pres* for present tense morpheme, *dec* for declarative morpheme, and *fut* for future tense morpheme.

in Pusan before coming to the U.S., and two of them had lived in Pusan over 20 years and lived in Seoul for 8 ~12 years before coming to the U.S.

#### *2.1.1.3 Procedures*

The test sentences were presented in standard Korean orthography on a laptop, in a timed PowerPoint presentation. There was a five-second interval between stimuli. Subjects read each presented sentence aloud ten times in succession, and then went on to the next sentence. For example, the subjects read the sentence for "M in unsuffixed words" ten times first, and then read another sentence for "M in monosyllabic suffixed words" ten times. The typical way to present the test sentences is to present them to the subjects in a random order, in order to avoid an order effect. However, it was practically problematic in this experiment, because the test words are tone minimal pairs and the test sentences for each tone type consist of almost same words. Therefore, I found in my pilot study that the test sentences in a random order were confusing to some subjects, and thus some of them had difficulties in producing the test sentences in a consistent way with a timed presentation. For this reason, the stimuli were not shuffled in this production experiment.

The subjects were told to keep their loudness consistent, and if they failed to produce a sentence with the same loudness, they were asked to read the sentence again, because the F0 peaks are higher in louder speech than in less emphatic speech (Liberman & Pierrehumbert 1984). I did not correct other mistakes or irregularities but the subjects could initiate to read any sentence again if they wanted to.

The recordings were made directly into a personal laptop using a Sennheiser headset mike in a sound-treated room at the Phonetics lab of the Linguistics department, University of Texas at Austin.

#### *2.1.1.4 Measurements*

The recordings then were digitized and analyzed using the PCquirer software package from SciConRD. For each sentence, synchronized displays of the sound waveform, a wide-band spectrogram, and an F0 track were produced.

Specifically, the following time points and F0 values were measured (*see* Fig. 2.2 for an illustration):

- *Syllable boundaries*: they were measured in the sound waveform expanded so that each individual pulse was clearly distinguishable. Since the test syllable is the first syllable and begins with an utterance-initial [m], the *onset of the test syllable* was measured at the beginning of the amplitude in the waveform and this point was confirmed by comparison with the spectrogram. It is illustrated by (a) the first vertical line in Fig. 2.2. Since the test syllable ends with the onset of the tap at the onset of the next syllable, the *offset of the test syllable* was measured at the onset of the marked drop in amplitude in the waveform and a fall to a minimal value of F1 in the spectrogram. It is indicated by (b) the second vertical line in Fig. 2.2.
- *Low F0*: it was measured at the point at which the local F0 value starts to rise during the test syllable (first syllable), indicated by the first arrow in Fig. 2.2. Both a time point and an F0 value at this point were measured.
- *F0 peak onset*: it was measured at the highest point. If there is a series of equally or almost equally (varied within 2 Hz) high points, the first in the series was selected, indicated by the second arrow in Fig. 2.2. Both a time point and an F0 value at this point were measured.

- *F0 peak offset*: it was measured at the first F0 value to be followed by a string of successively (at least four points) lower values was selected, indicated by the third arrow in Fig. 2.2. Both a time point and an F0 value at this point were measured.

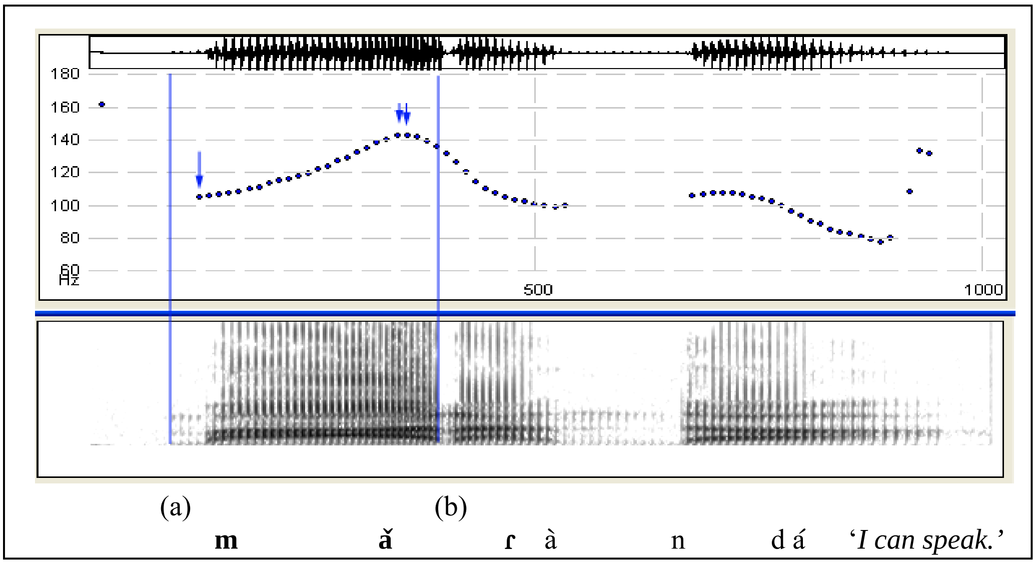


Figure 2.2: Sample display with five measurement points

Such time points and F0 values are measured to calculate the following dimensions in (11):

(11) Measurements

a. the test syllable duration: the first syllable offset - the first syllable onset

b. F0 peak delay: F0 peak onset - the first syllable onset

c. F0 fall delay: F0 peak offset - the first syllable onset

d. peak plateau: F0 peak offset- F0 peak onset

e. The low F0 and peak F0 values

### 2.1.2 Results

#### *2.1.2.1 The timing of F0 peak*

*2.1.2.1.1 Relative peak delay*

The difference in the timing of F0 peak between tone class we call M-class and R-class is first considered. The results show that the F0 peak came later in R-class items than in M-class items, across all suffix types.

Fig. 2.3 presents typical F0 tracks for the two tone types in three suffix types. Fig. 2.3 (a,b,c) shows M tone types and Fig. 2.3 (d,e,f) R tone types. The first vertical line in each display indicates the onset of the first syllable, the second one marks the offset of the first syllable. The test syllable is the first syllable for all tone types. The arrow indicates the peak F0. The test words are indicated in bold, and the suffixes are italicized.

It can be seen in this figure that the F0 peak comes *in* the first syllable for M-class item across suffix types and for R-class item in unsuffixed words, but occurs *after* the first syllable R-class item in suffixed words.

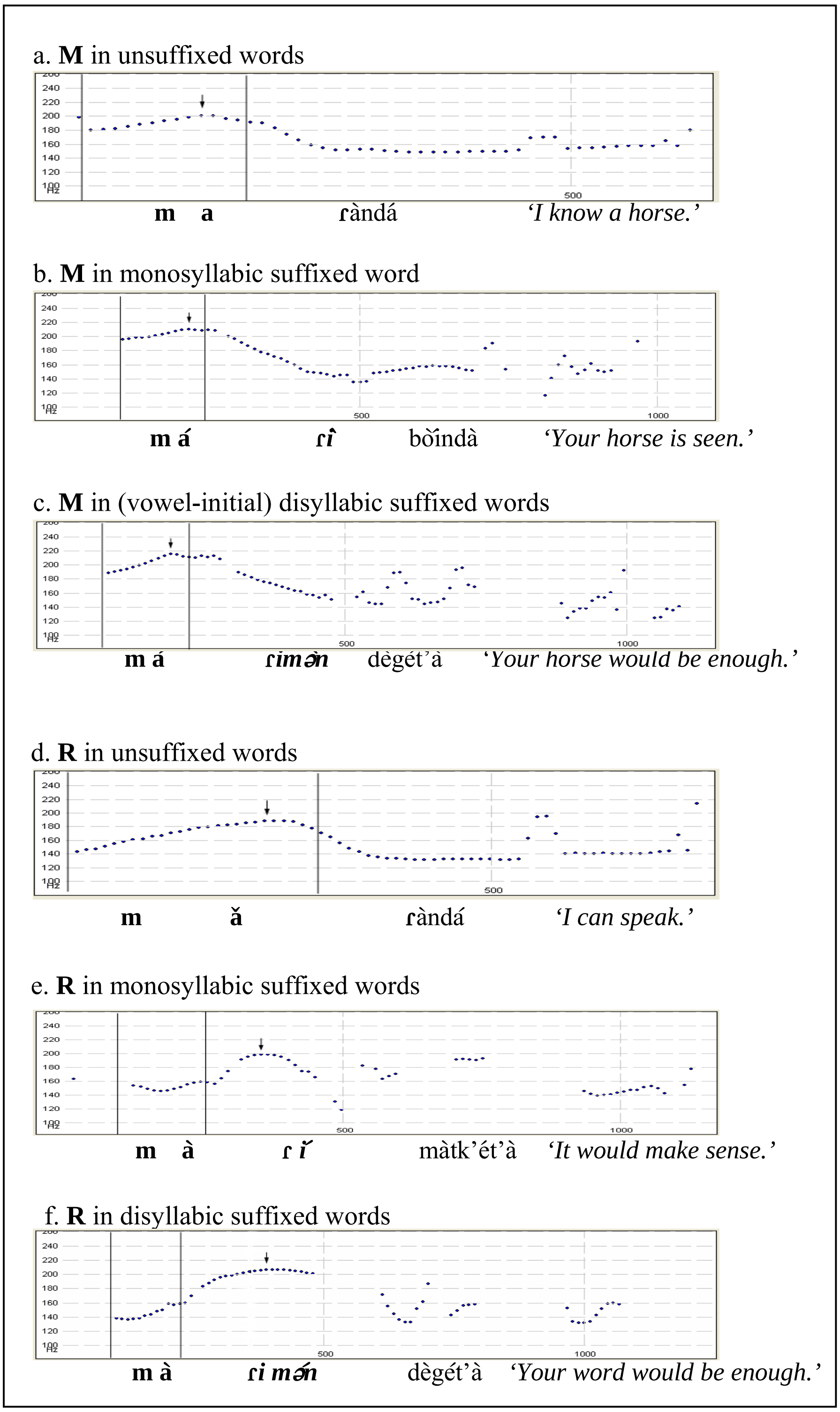



Figure 2.3: Typical effects of M and R on the timing of F0

A quantitative measure of the F0 peak timing difference is *relative peak delay* (the peak delay divided by the test syllable (i.e., the first syllable) duration); it provides the percentage of the test syllable duration at which the F0 peak is first attained. Fig. 2.4 shows the variation in relative peak delay in a box plot graph, with tone type and suffix type compared for each speaker. It shows a 5-point summary ($90^{th}$, $75^{th}$, $50^{th}$ (median), $25^{th}$, and $10^{th}$ percentiles) with five horizontal lines for each indicated group.

The relative peak delay "1", marked by a horizontal line, indicates that the peak comes precisely at the offset of the test syllable, "less than 1" indicates that the peak comes *before* the end of the test syllable (i.e., it comes *during* the test syllable), and "greater than 1" indicates that the peak comes *after* the test syllable. The empty boxes mark M, and dark boxes mark R.

As seen in Fig. 2.4, for all six speakers, the median relative peak delay is greater for dark boxes than for empty boxes for each suffix type, indicating that the peak delay is greater for R than for M for suffix type. The median relative peak delay was "greater than 1" for R in suffixed words, indicating that a peak came *after* the first syllable in suffixed R. The median relative peak delay was "less than 1" in M. This shows that a peak came *in* the first syllable for M across suffix types.

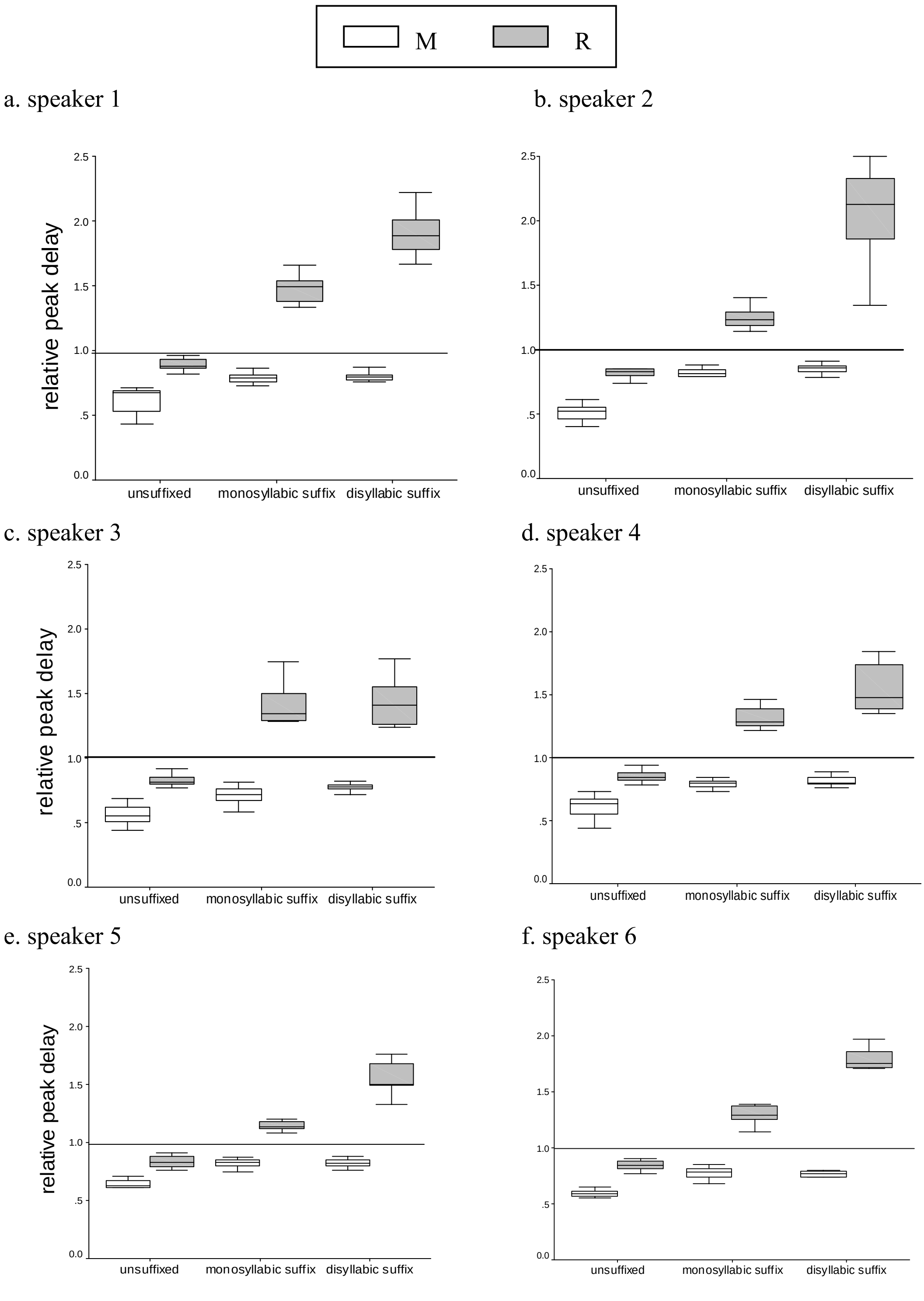



Figure 2.4: *Relative peak delay* of M and R by tone type and suffix type

Within M, the median is lower in the unsuffixed words than in the suffixed words. This indicates that the peak came earlier in the syllable in unsuffixed (word-final syllable) than in suffixed. Within R, the median is higher in disyllabic suffixed words than in monosyllabic suffixed words, and is higher in monosyllabic suffixed words than in unsuffixed words. The median is "greater than 1" for R in suffixed words, indicating that the peak was *after* the first syllable for R.

The statistical analysis was conducted. There is a continuous dependent variable, i.e., *relative peak delay*, and two categorical independent variables, i.e., *tone* types (M and R) and *suffix* type (unsuffixed words, monosyllabic suffixed words, and disyllabic suffixed words). The data was also gathered using repeated measurements from 6 different subjects. Therefore, a repeated-measured factorial ANOVA with subject as the between-subject factor, and tone type and suffix type as within-subject factors was conducted.

There was a significant main effect of *tone* type ($F(1, 54)=2519$, $p < .001$), indicating that mean relative peak delay is significantly greater in the R-class than in the M-class. This result support the hypothesis that the peak delay is greater for R-class than for M-class across the suffix type. The main effect of *suffix* type was also significant ($F(1,54)=1276$, $p < .001$), indicating that mean relative peak delay is significantly greater in suffixed words than in unsuffixed words. The *interaction* of tone type and suffix type was also significant ($F(1,54)=474$, $p<.001$). This might be because the F0 peak is in the second or third syllables for suffixed R-class while the peak occurs in the initial syllable for suffixed M-class.

We also need to test the hypothesis that the peak comes later in R with a disyllabic suffix than in R with a monosyllabic suffix. Post-hoc tests using Fisher's PLSD were conducted to evaluate pair-wise comparisons among the means. The significant differences were found among six groups as follows; “R, disyllabic suffixed” > “R, monosyllabic suffixed” > “R, unsuffixed” > “M, disyllabic suffixed”, “M, monosyllabic suffixed” > “M, unsuffixed”. This supports the hypothesis that the peak comes later in R with a disyllabic suffix than in R with a monosyllabic suffix.

Within M-class items, the peak was earlier in unsuffixed words than in suffixed words, and this suggests the earlier peak delay in a final position that has been observed in other languages (English [Steel 1986, Silverman & Pierrehumbert 1990], Spanish [Prieto et al. 1995], Kinyarwanda [Myers 2003]).

Overall, the result for the relative peak delay shows that the peak was later for R-class than for M-class within the syllable in unsuffixed words. In suffixed words, it was shown that the peak occurs *in* the first syllable for M across the suffix type. In comparison, the peak occurs *after* the first syllable for suffixed R-class. These findings are consistent with the proposed transcriptions in (9) that the high tone occurs in the initial syllable for M-class in suffixed words.

However, although it was seen that the peak was *after* the first syllable for R with a suffix, it was not determined whether it is in the second or third syllable as suggested in (9), because the test syllable was the first syllable in this analysis. We can see whether the peak is in the second syllable or third syllable for R in suffixed words, by considering the relative peak delay in relation to the *second* syllable. If the relative peak delay is “less than 1”, it means that the peak was in the *second* syllable. If it is “greater than 1”, it indicates that the peak was in the *third* syllable.

Fig. 2.5 shows the variation in relative peak delay of R-class relative to the second syllable in a box plot graph, with suffix types compared for each speaker.

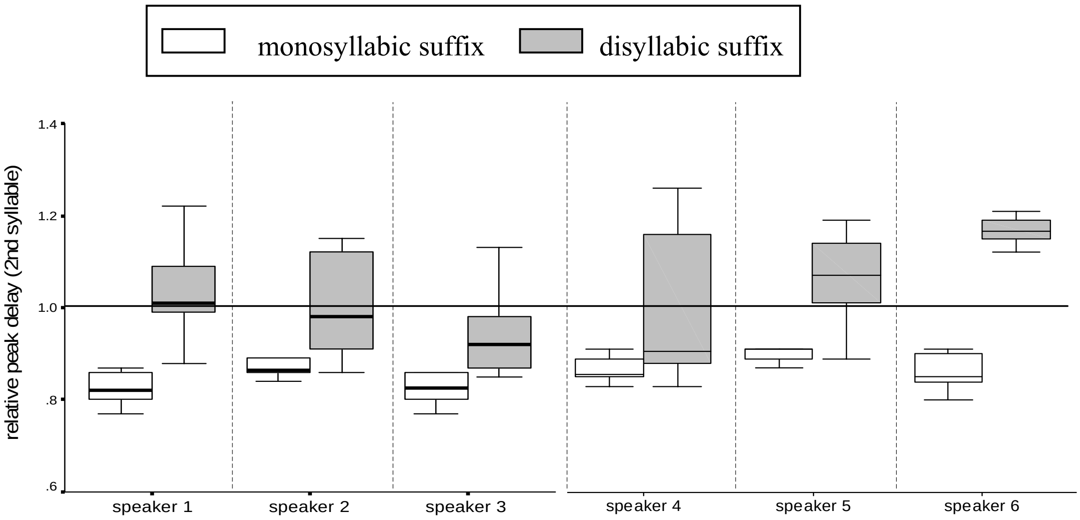


Figure 2.5: *Relative peak delay* relative to the second syllable for R

It is seen in this figure that the peak was in the second syllable for monosyllabic suffixed words, marked by empty boxes, for all speakers. However, there is great variation for disyllabic suffixed words, both between speakers and within speakers, indicating that the peak was sometimes in the final portion of the second syllable, but sometimes in the early portion of the third syllable. However, the consistent pattern for all six speakers is that the peak is later in disyllabic suffixed words than in monosyllabic suffixed words.

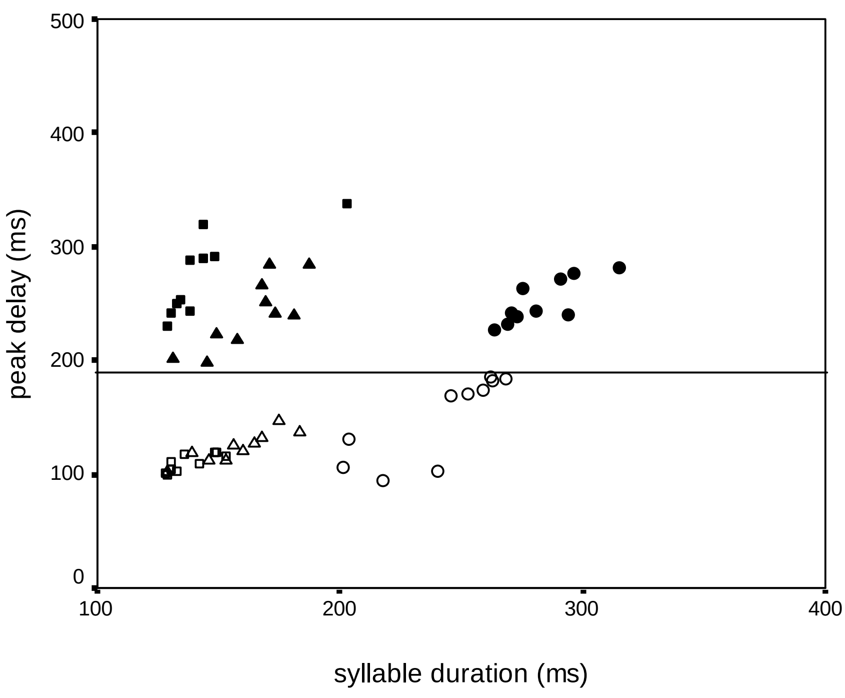
*2.1.2.1.2 Syllable duration*

The relative peak delay depends partly on peak delay (the interval from the onset of the syllable to the onset of the F0 peak (ms)) and partly on syllable duration, thus these two aspects need to be considered separately to see if the peak delay varies in the same way that the relative peak delay does, i.e., whether the difference in relative peak delay is due to peak delay differences or to test syllable (first syllable) duration differences. In addition, the hypothesis that the syllable duration is longer for R than for M needs to be tested.

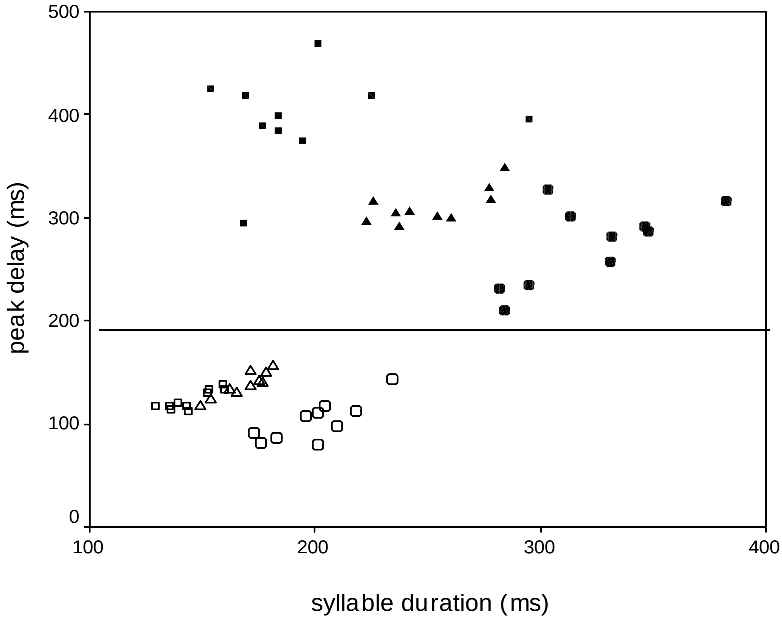
The data for peak delay and syllable duration are plotted in Fig. 2.6, where the x-axis represents the syllable duration and the y-axis the F0 peak delay. Datapoints from M are plotted with empty shapes, while those from R are plotted with dark ones. The "circles" mark the unsuffixed words, "triangles" mark the monosyllabic suffixed words, and "small squares" mark the disyllabic suffixed words.

| | | |
|---|---|---|
| ○ M, unsuffixed | △M, monosyllabic suffixed | □M, disyllabic suffixed |
| ● R, unsuffixed | ▲R, monosyllabic suffixed | ■R, disyllabic suffixed |

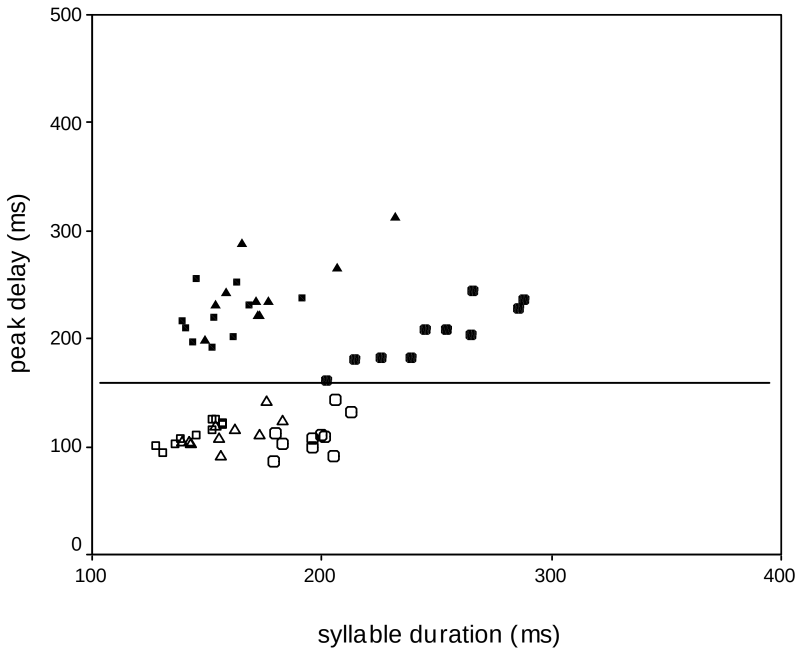
a. speaker 1



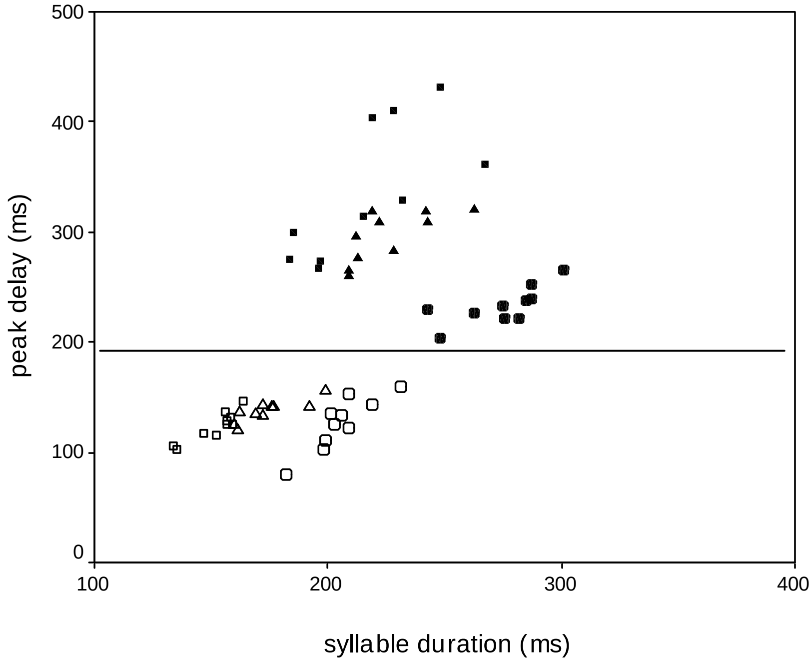
b. speaker 2



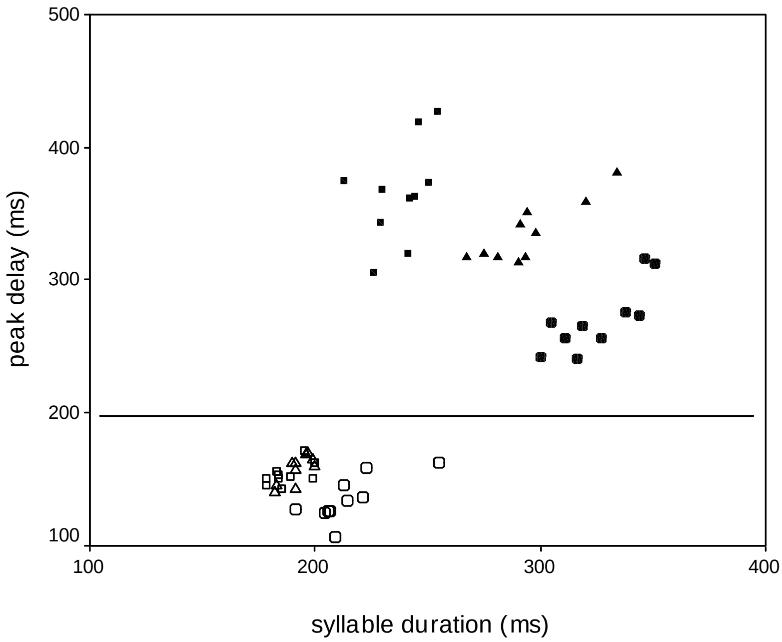

c. speaker 3



d. speaker 4




e. speaker 5



f. speaker 6

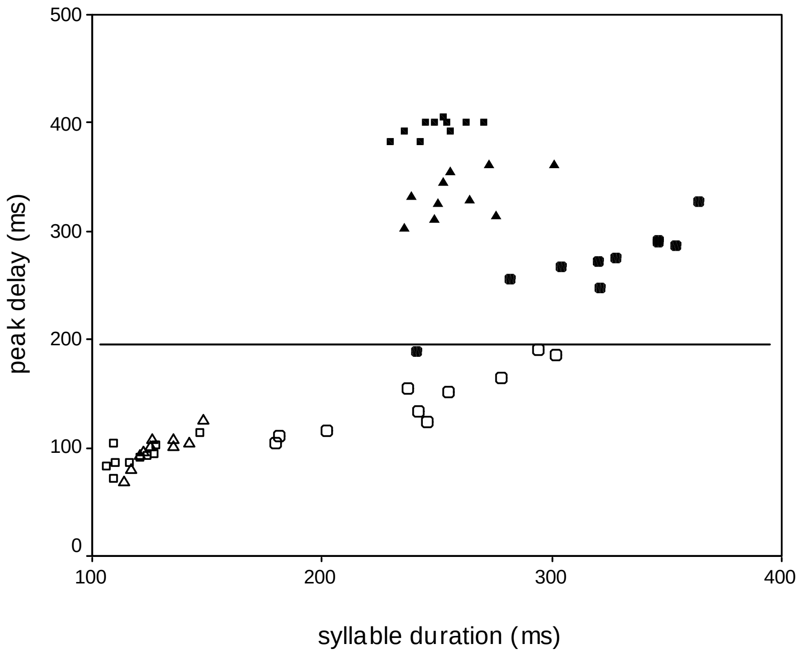



Figure 2.6: Peak delay relative to syllable duration for M and R by suffix type

The dark shapes in Fig. 2.6 are above the empty shapes, indicating a difference in absolute peak delay between M-class and R-class, as indicated by a horizontal line. Further, the dark shapes are grouped to the right of the empty ones *within* suffix group, indicating that the test syllable (the initial syllable) duration is longer for R-class than M-class within suffix type. In addition, the circles are grouped to the rightmost, and squares are grouped to the leftmost for each tone type, showing that the unsuffixed words are the longer than monosyllabic words, and monosyllabic words are longer than disyllabic words. This is because the suffix types differ in word length, and word length is known to affect segment duration, i.e., longer words have shorter syllables (Lindblom 1981).

Although there is a variation in how tightly the points within a group hug a linear trendline and they differ in the slop of that line between speakers, the points display a linear trend in general, with F0 peak delay being greater for longer syllables, as represented by the slope of distributions. The fact that the distributions are sloped rather than horizontal indicates that peak delay is not constant but varies systematically as a function of syllable duration. That is, as the duration increase, so does the peak delay.

In order to test the hypothesis that the syllable duration is longer for R-class than for M-class, the data of "*syllable duration*" were analyzed statistically by means of repeated measures ANOVA in which syllable durations were dependent measures. Subject was the between-subjects factor, and tone type (M and R) and suffix type (unsuffixed, monosyllabic suffixed, disyllabic suffixed) were the within-subjects factors.

There was a significant main effect of *tone* type, $F(1, 54)=767$, $p <.001$, indicating that mean syllable duration is significantly longer in R-class than in M-class. The main effect of *suffix* type was also significant ($F(1,54)=878$, $p <.001$), showing that the mean is significantly longer for unsuffixed words than for suffixed words. The *interaction* of tone type and suffix type was significant ($F(1,54)=24$, $p<.001$), with a greater unsuffixed word effect for R. This might be because the final syllable lengthening effect in unsuffixed words and intrinsic long syllable duration of R-class are combined in unsuffixed R-class. Post-hoc tests using Fisher's PLSD were conducted to evaluate pair-wise comparisons among the means. It showed the following significant differences among six groups; "R, unsuffixed" > "R, monosyllabic suffixed" > "M, unsuffixed" > "R, disyllabic suffixed" > "M, monosyllabic suffixed" > "M, disyllabic suffixed".

This result supported the hypothesis that the syllable duration is longer for R-class than for M-class in unsuffixed words. However, the result also found that the test syllable (initial syllable) duration is longer for R-class than for M-class in suffixed words as well. We expected that the syllable is longer for R-class than for M-class in unsuffixed words, because the R-class might have two tone targets (low pitch + peak pitch), and this would cause the longer syllable duration for R-class in unsuffixed words. The fact that the syllable duration was longer for R than for M in suffixed words suggests that the initial syllable of R-class in suffixed words is also a rising tone than simply a level low tone. Although the peak is undershoot in the first syllable and is realized *after* the first syllable, the nature of the tone on the initial syllable might be the part of the rising tone, i.e., the initial low tone target of the rising tone.

The syllable duration was also significantly longer in unsuffixed words for each tone type, and it appears that there is a strong final lengthening effect (Oller 1973, Klatt 1976, Lindblom et al. 1981). This also might be responsible for the lower relative peak delay in word-final syllable than in other syllables.

*2.1.2.1.3 Peak delay*

Next, the data of "*peak delay*" was analyzed with statically in order to see if the peak delay varies in the same way that the relative peak delay does. The repeated measures ANOVA test with a peak delay value as the dependent measure. Subjects were the between-subject factor, and tone type and suffix type were the within-subject factors.

The results found the significant main effects of *tone* type ($F(1,54) = 3235$, $p<.001$), indicating that the mea peak delay was significantly greater for R than for M. The main effect of *suffix* type was also significant ($F(1,54)= 97$, $p< .001$), indicating that the mean was significantly greater for suffixed words than for unsuffixed words. The main effect of *interaction* of tone and suffix was significant ($F(1,54)= 146$, $p< .001$), with a great suffix effect of R-class. This might be because the peak occurs in the second and the third syllable for suffixed R, while the peak occurs in the first syllable for M across suffix type. Post-hoc tests using Fisher's PLSD showed the following significant differences among six groups; “R, disyllabic suffixed” > “R, monosyllabic suffixed” > “R, unsuffixed” > “M, unsuffixed”, “M, monosyllabic suffixed” > “M, disyllabic suffixed”.

This result showed that although the *relative* peak delay was shorter for M in unsuffixed words than for M in suffixed words, the *absolute* peak delay was not significantly shorter in unsuffixed M-class than in suffixed M-class. Rather, it was significantly longer in unsuffixed M-class than in disyllabic suffixed M-class.

The result of syllable duration and peak delay can be interpreted that the fact that relative peak delay was significantly less in unsuffixed words was due more to the lengthening of the word-final syllable than to the peak delay difference. The extra duration of these final syllables did not affect peak delay, but did affect relative peak delay by increasing the denominator (test syllable duration) that peak delay is divided by when calculating relative peak delay. Therefore, when the syllable duration is lengthened due to the prosodic factors, such as syllable position, the peak seems to move close to the syllable onset. However, it appears to be due to the syllable lengthening effect because the absolute peak delay is increased as the syllable duration is lengthened. In contrast, when the syllable duration is lengthened due to phonetic factors, such as in R, it was seen that the F0 peak moves away from the syllable onset along with the syllable offset.

To summarize, the analysis of the timing of F0 peak showed that there are differences between M and R, i.e., the peak was later in the syllable for R than for M in unsuffixed words. The peak was earlier in unsuffixed words (word-final) than in suffixed words (non-final), but this appears to have been due more to the effects of syllable duration than to a difference in how the F0 peak was timed with respect to the syllable. In suffixed words, the peak was in the *first* syllable for M both with a monosyllabic suffix and a disyllabic suffix. The peak was in the *second* syllable for R with a monosyllabic suffix, and in the *third* syllable for R with a disyllabic suffix, as suggested in (9).

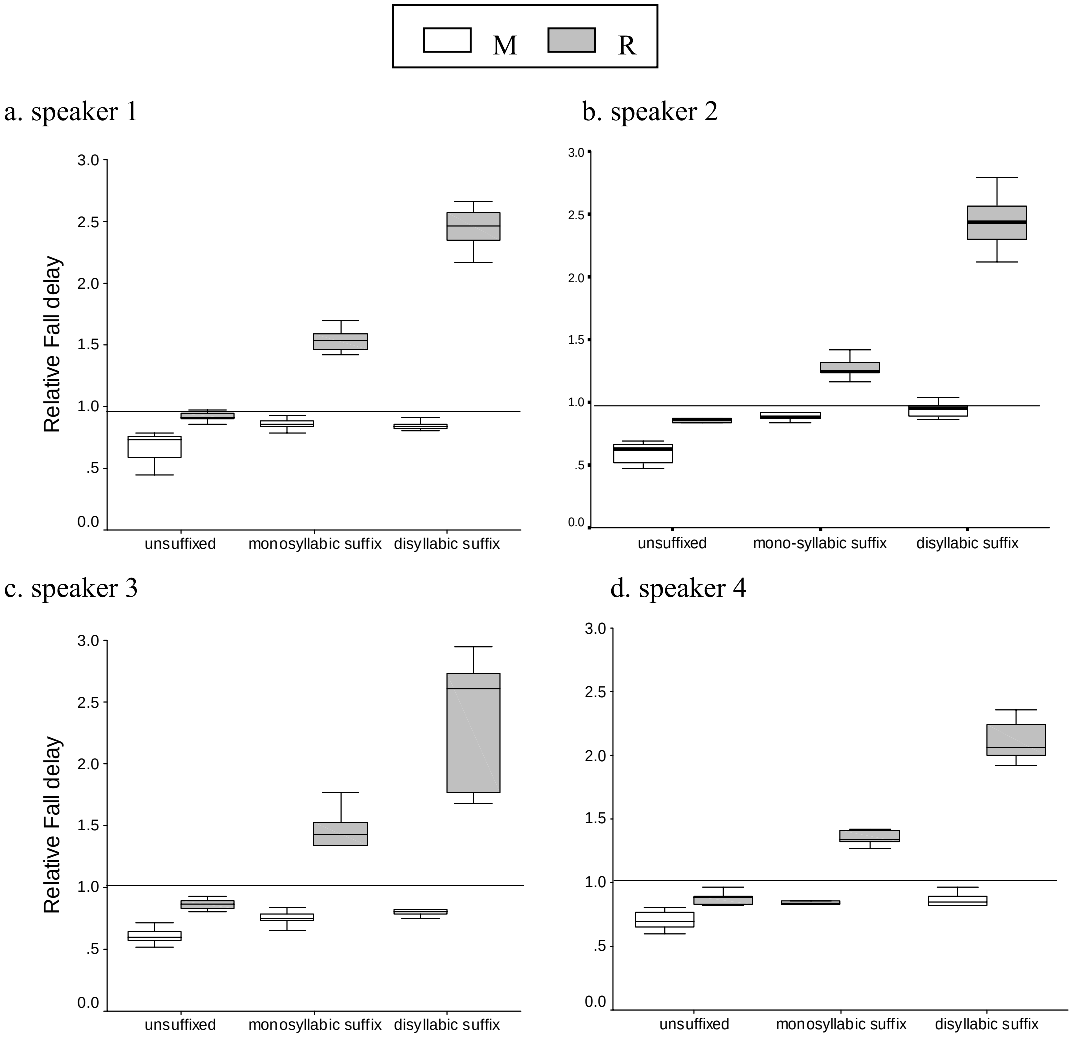
*2.1.2.2 The timing of F0 fall*

The difference in the timing of F0 fall between M-class and R-class is considered in this section. This measurement shows when the F0 fall comes for each tone type. We've already seen in the preceding section that F0 peak is in the first syllable for M in all suffix type, and the peak is in the second syllable for R in monosyllabic suffixed words and in the early portion of the third syllable for R in disyllabic suffixed words.

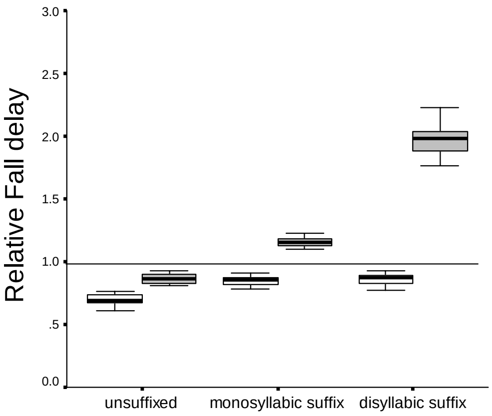
In this section, we test whether the F0 fall is also in the same syllable with the F0 peak. For example, if the high tone is associated with both the second syllable and the third syllable for R in disyllabic suffixed words as suggested by transcription-based data (Kim 1996), then the peak would be in the second syllable, but the fall would be in the third syllable. However, the preliminary work found that the peak usually occurs only in the third syllable although the second syllable also has quite high F0 values. Hence, we can hypothesize that the F0 fall would be in the same syllable with the F0 peak, and the pattern of the timing of F0 fall would be similar with the timing of F0 peak.

The specific hypotheses were that the F0 fall is later for R than for M across suffix type, and that the F0 fall is later in disyllabic suffixed words than in monosyllabic suffixed words for R. The result supported the hypotheses. A quantitative measure of the difference in the timing of F0 fall is the *relative fall delay* (the fall delay divided by the test syllable duration); it provides the percentage of the syllable duration at which the F0 fall is first attained. The important result about relative fall delay is that it mirrors the pattern of relative peak delay.

Fig. 2.7 shows the variation in relative fall delay in a box plot graph with tone type and suffix type compared for each speaker. It is seen that the medians in the empty boxes marking M are "less than 1" for all speakers, indicating that the F0 fall came in the first syllable in M across suffix type. The medians in the dark boxes marking R are “greater than 1” for suffixed words, indicating that the F0 fall came after the first syllable for H in suffixed words. It shows that the relative fall delay is greater in R than in H for each suffix type for all speakers.



e. speaker 5




f. speaker 6

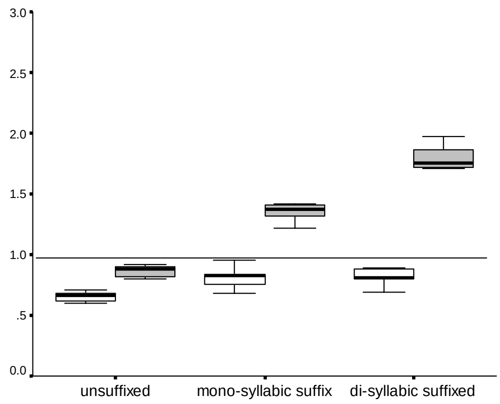



Figure 2.7: The *relative fall delay* of M and R by tone and suffix type

The repeated measures ANOVA test with a subject as the between-subject factor, and tone type (M and R) and suffix type (unsuffixed, monosyllabic suffixed, disyllabic suffixed) as the within-subject factors was conducted. The relative fall delay was the dependent variable.

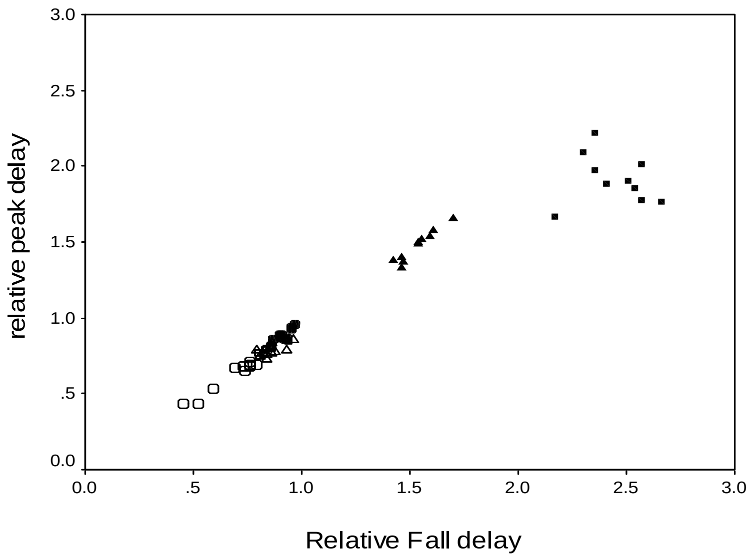
There was a significant main effect of *tone* type (F (1, 54)=2532, p <.001), indicating the mean relative fall delay was significantly greater for R than for M. This result supports the hypothesis that the F0 fall delay is greater for the tone class we call R-class than for tone class we call M-class across suffix type. The main effect of *suffix* type was also significant (F(1,54)=1428, p <.001), showing the means were significantly greater for suffixed words than for unsuffixed words. It is the interaction of tone and suffix that is significant in this experiment, since two combinations of tone and suffix (i.e., R with a monosyllabic suffix, and R with a disyllabic suffix) have much greater means than other combinations. The *interaction* of tone type and suffix type was significant (F (1,54)=747, p<.001), with a greater suffixed word effect for R-class. This is because the F0 fall is in the second or third syllables in suffixed words for R. Post-hoc tests using

Fisher's PLSD among six groups supported the hypothesis that the F0 fall delay is greater for disyllabic suffixed R-class for monosyllabic suffixed R-class.

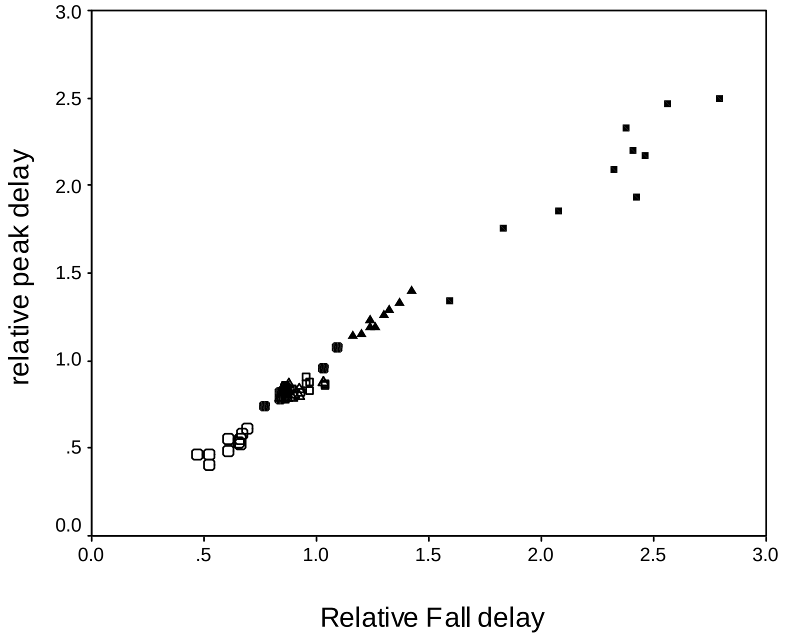
The important aspect of relative fall delay is that it is quite similar in pattern to relative peak delay, as noted earlier. If we were to plot the relationship between the relative peak delay (on the y-axis) and the relative fall delay (on the x-axis), we might see the results shown in Fig. 2.8. Datapoints from M are plotted with empty shapes, while those from R are plotted with dark ones. The "circles" mark the unsuffixed words, "triangles" mark the monosyllabic suffixed words, and "small squares" mark the disyllabic suffixed words. From the graph we can see a linear relationship, i.e., as the relative peak delay increases, so does the relative fall delay. Thus this scattergram demonstrates a *positive correlation* between the relative peak delay and relative fall delay.

| | | |
|---|---|---|
| ○M, unsuffixed | △ M, monosyllabic suffixed | □M, disyllabic suffixed |
| ● R, unsuffixed | ▲ R, monosyllabic suffixed | ■R, disyllabic suffixed |

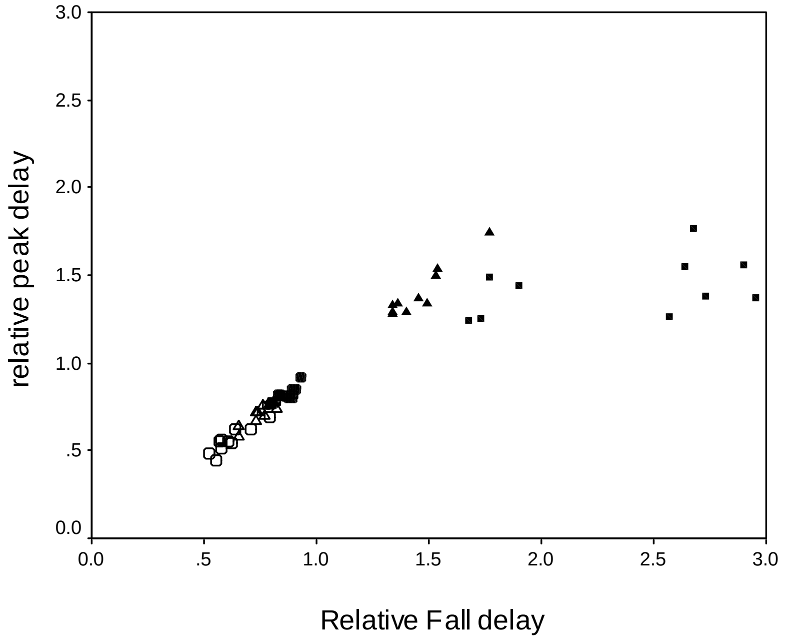
a. speaker 1



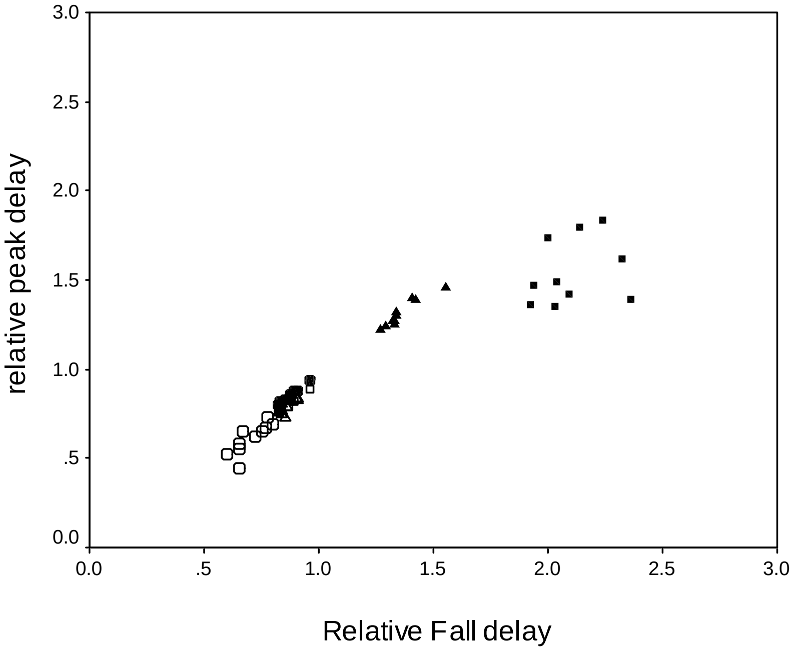
b. speaker 2



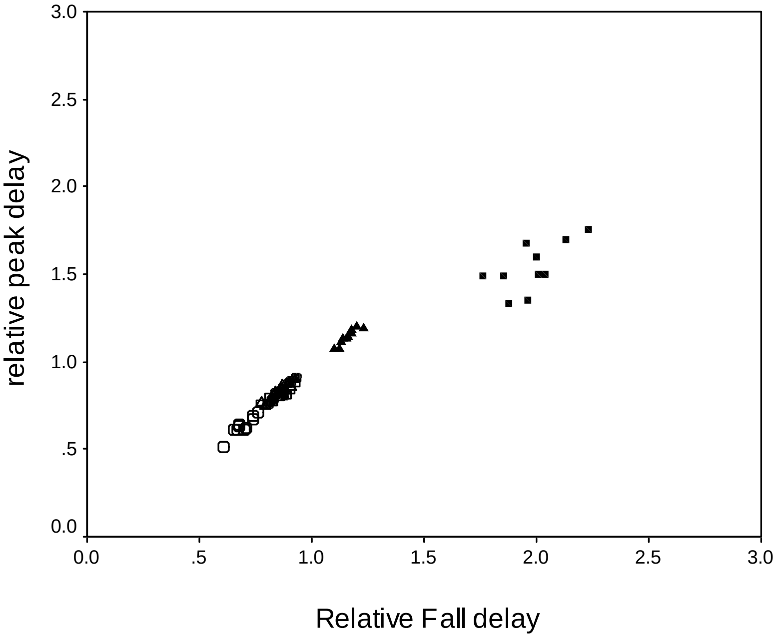

c. speaker 3



d. speaker 4




e. speaker 5



f. speaker 6

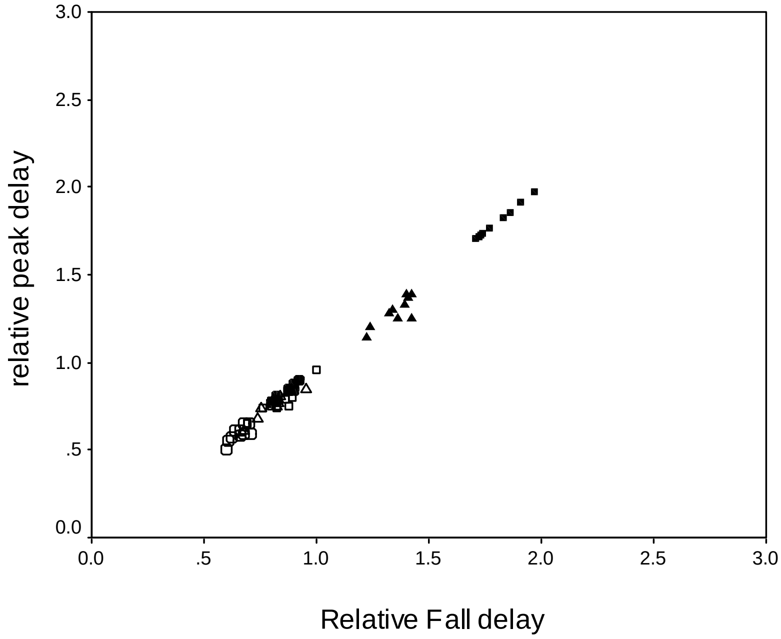
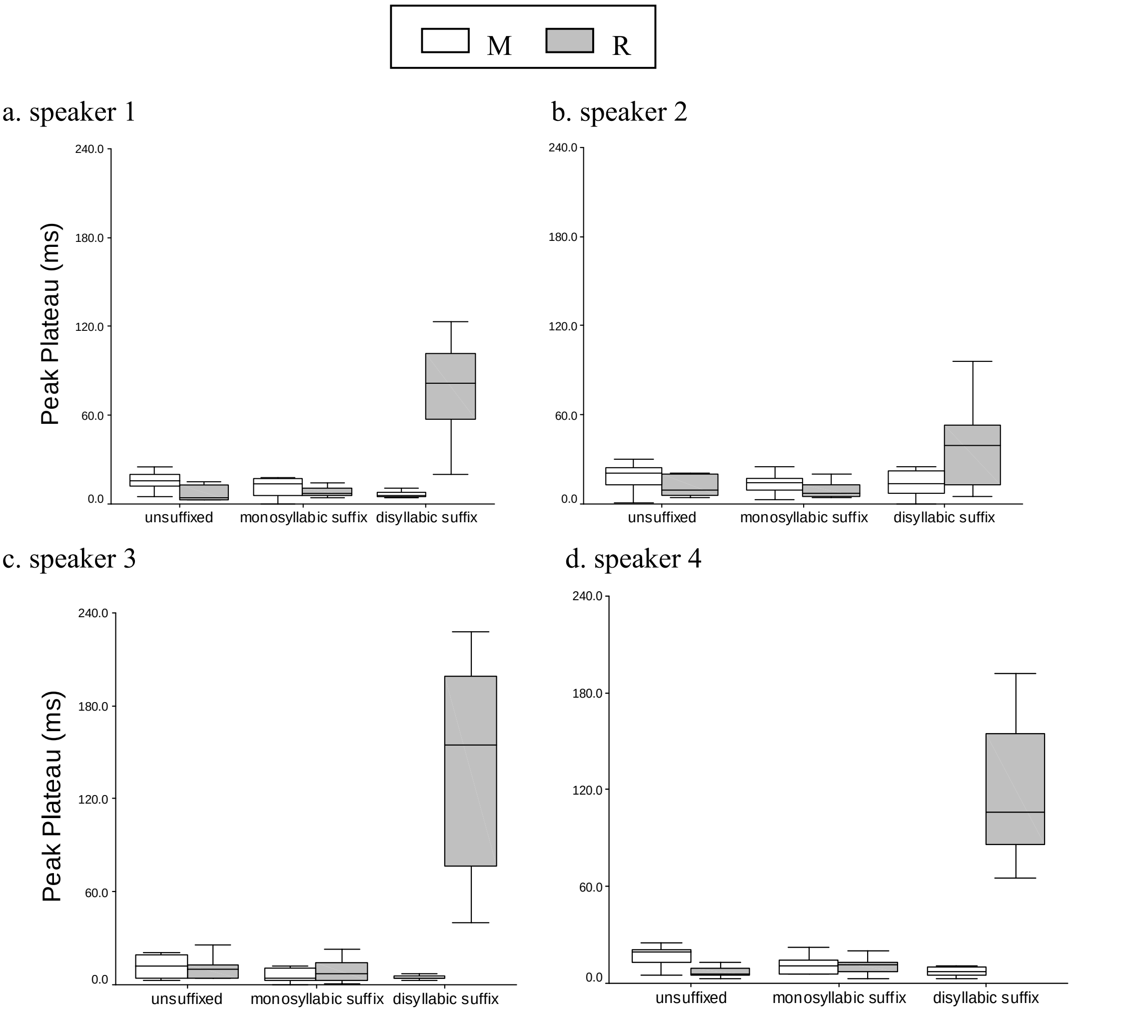



Figure. 2.8: Relative peak delay against relative fall delay

On the whole, the data for the timing of the F0 fall supported that hypotheses that the F0 fall delay is greater for the tone class we call R-class than for the tone class we call M-class and it is greater for the disyllabic suffixed R-class than for the monosyllabic suffixed R-class in a similar way with the peak delay.

### *2.1.2.3 Peak plateau*

This section tests whether the peak plateau (the onset of the F0 fall - the onset of the F0 peak) is longer for the disyllabic suffixed words than the others. If the F0 peaks are stretched over two syllables in disyllabic suffixed R-class as suggested in Kim (1996), then their peak plateaus would be longer than the others. The Fig. 2.9 shows the variation in *peak plateau* in a box plot graph with tone type and suffix factor for each speaker. The peak plateau is longer for disyllabic suffixed R than the others, and it appears to support the claim that the high tone is associated with two syllables (the second syllable and the third syllable). However, the notable pattern is that the variation range of the peak plateau for disyllabic suffixed R is considerably large, both between subjects and within subjects. That is, the difference of the 90th percentile and the 10th is great, indicating that the peak plateau was sometimes very long and sometimes very short and thus the pattern is not obligatory or regular even within subjects.



e. speaker 5 f. speaker 6

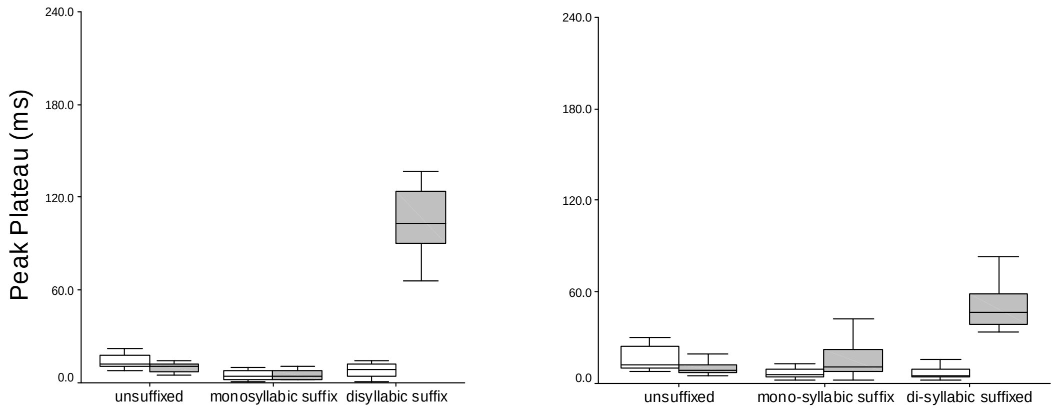



Figure 2.9: The *peak plateau* of M and R by tone and suffix type

The repeated measures ANOVA was used for a statistical test. There were two within-subject factors: tone type (M and R) and suffix type (unsuffixed, monosyllabic suffixed, and disyllabic suffixed). Subjects were considered as a between-subject factor. The *peak plateau* was the dependent variable.

Although there were significant main effects of *tone type* ($F(1,54)=193$, $p < .001$) and *suffix type* ($F(1,54)=170$, $p < .001$), these main effects are probably significant only due to the interaction of tone type and suffix type, because peak plateau is longer in R than in M, not across the board, but only in disyllabic suffixed words. The *interaction* of the tone type and suffix type was significant ($F(1,54)=289$, $p < .001$), with great suffixed word effects for R. This might be because both second and third syllables have high F0 ranges and thus peak sometimes consisted of a single point in the third syllable, but sometimes was held for some time from the second syllable to the third syllable, yielding long peak plateaus. A pair-wise comparison by a post-hoc test showed that the mean peak plateau is significantly greater for “R, disyllabic suffixed” than for “M, unsuffixed”, and greater for “M, unsuffixed” than for the other groups.

This result shows that the peak plateau was longer for the tone class in question, i.e., “R, disyllabic suffixed” than the others, and it seems to support the claim that the high tone is associated with two syllables (Kim 1996). However, it was already seen in earlier sections that the peak onset and offset were usually occur in the same syllable rather than the peaks are stretched over two syllables. In addition, the within-subject variation in peak plateau was great as seen in Fig. 2.8 above. It suggests that the long peak plateau pattern is not consistent or obligatory. One might wonder why the peak plateau is such longer in disyllabic suffixed R-class. This might be because the high tone is realized with gradual F0 rising through the second syllable up to a peak near the onset of the third syllable, results in the quite long peak plateau around the onset of the third syllable.

*2.1.2.4 The low F0 value*

The preliminary observation found that the initial pitch is lower for the tone class we call R than for the tone class we call M, and thus the hypothesis was that the low F0 is lower for R than for M. The lowest F0 point usually occurs at the syllable onset for M in all suffix types, but it occurs a little later for R, i.e., around the onset of the vowel, in all suffix types.

Fig. 2.10 shows the variation in *low F0* value (Hz) in a box plot graph with tone type and suffix factor for each speaker. It is seen in this figure that the median is generally lower in dark boxes than in empty boxes for each suffix type, indicating that the low F0 is lower in R than in M for each suffix type.

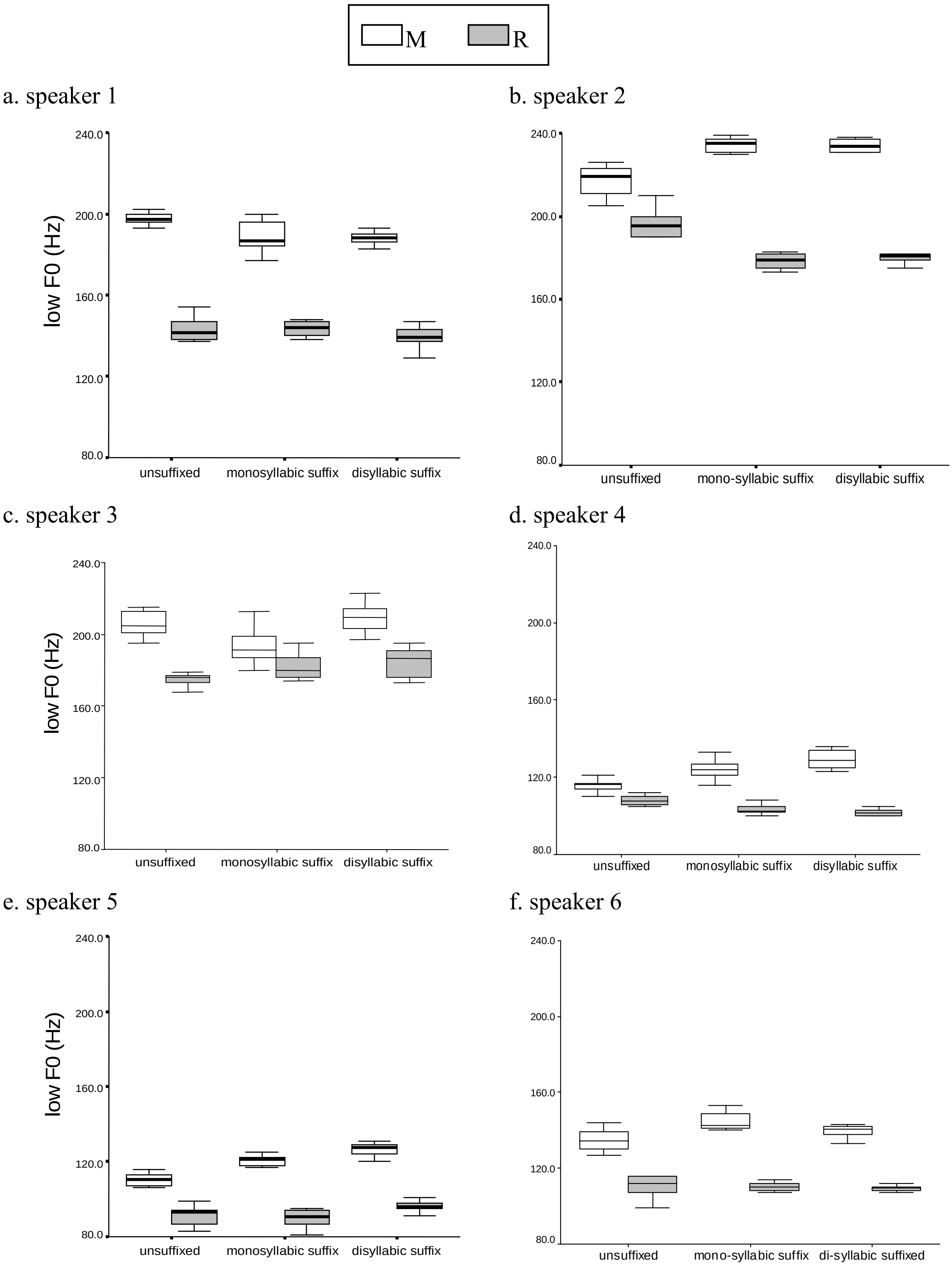



Figure 2.10: The *low F0* of M and R by tone and suffix type

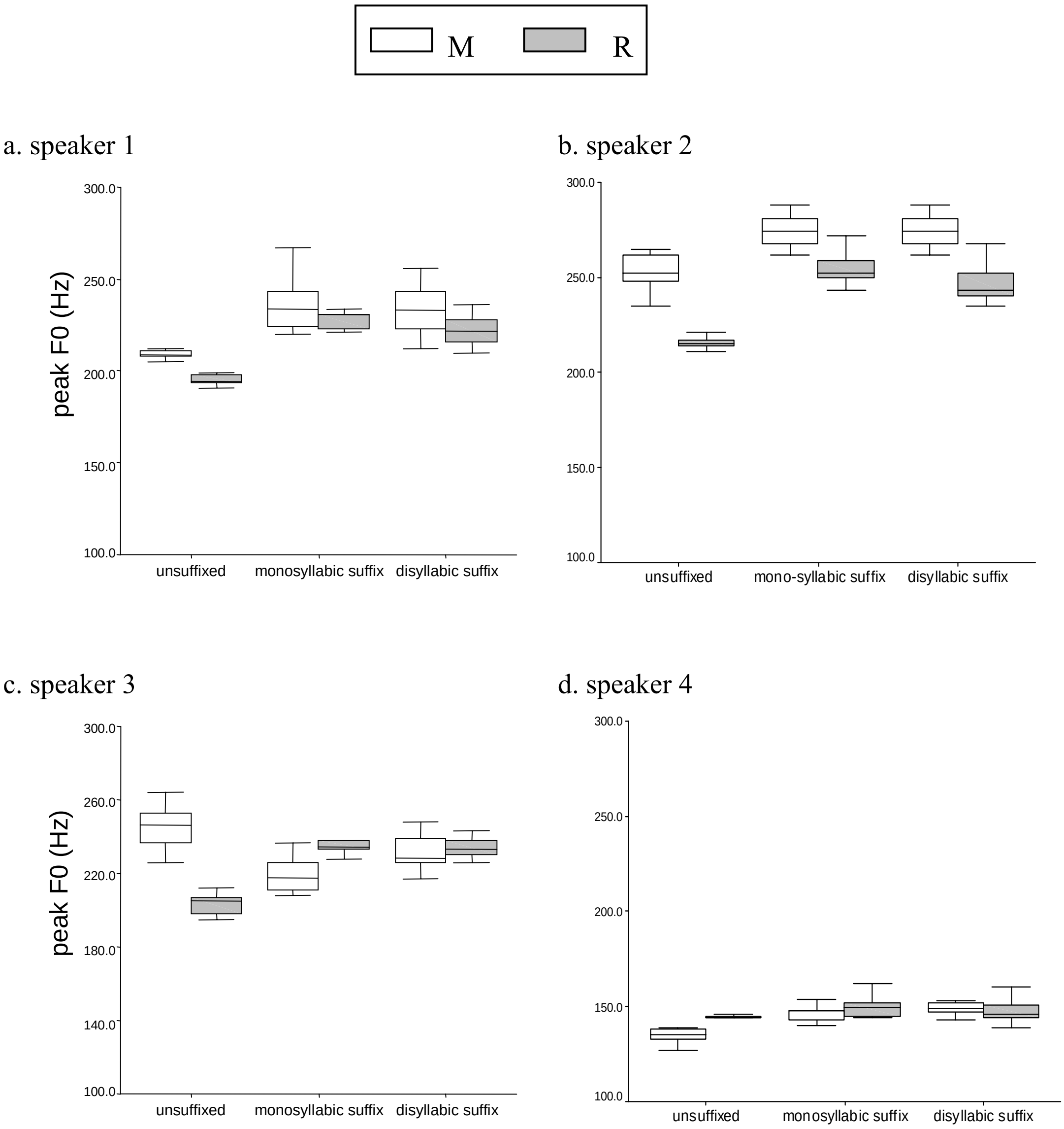
For analysis of initial F0, the pitch values in Hz were converted into z-scores in order to compare across speakers with radically different F0 ranges (e.g., men and women). The repeated measures ANOVA was performed: the dependent variable was the z-scores converted from the low F0 values in Hz. The tone type (M and R) and suffix type (unsuffixed, monosyllabic suffixed, disyllabic suffixed) were tested as within-subject factors, subject were tested as between-subject factors.

There was a significant main effect of *tone* type ($F(1,54)=1643$, $p < .001$), indicating that the mean low F0 values (Hz) are significantly greater for M-class than for R-class. This supports the hypothesis that the low F0 is lower for R-class than for M-class across the suffix type. The *interaction* of the tone type and suffix type was also significant ($F(1,54)=71$, $p < .001$), with great suffix effects for M-class. This might be because the low F0 is higher for suffixed M-class than for unsuffixed M-class.

The Post-hoc tests using Fisher's PLSD showed the following significant differences among six: "M, disyllabic suffixed" > "M, monosyllabic suffixed" > "M, unsuffixed" > "R, disyllabic suffixed", "R, monosyllabic suffixed", "R, unsuffixed". This result suggests that the low point is quite stable and solid for R-class irrespective of suffix type, while it varies depending on the suffix type for M-class. More specifically, the low F0 value was significantly lower for unsuffixed M-class. This can be interpreted that the low F0 value is a target tone for R-class but it is a simply initial F0 value for M-class.

*2.1.2.5 The peak F0 value*

The hypothesis was that the peak F0 is also lower for the tone class we call R-class than for the tone class we call M-class due to the influence of the low initial F0 pitch of R-class. Fig. 2.11 shows the variations in F0 peak value in a box plot graph with tone type and suffix type compared for each speaker.



e. speaker 5 f. speaker 6

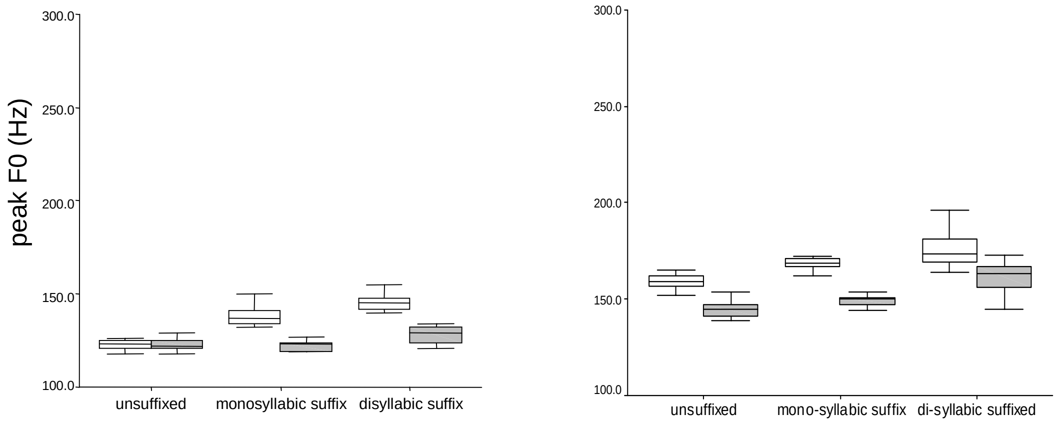



Figure 2.11: The *peak F0* by tone type (M and R) and suffix type

It is seen that the medians of the unsuffixed words are higher in empty boxes than in dark boxes for four speakers (speaker 1, 2, 3, and 6), indicating that the peak F0 is lower for R-class than for M-class in unsuffixed words for these 4 speakers. However, the median in unsuffixed words was greater for R-class than for M-class for one speaker (speaker 4), and there's no difference between two tone types in unsuffixed words for another speaker (speaker 5). It is also shown that the medians for M-class are higher in suffixed words than in unsuffixed words (except for speaker 3). Similarly, the medians for R-class are higher in suffixed words than in unsuffixed words (except for speaker 5).

The repeated measures ANOVA was performed: the dependent variable was the z-scores converted from the peak F0 values in Hz. The tone type (M and R) and suffix type (unsuffixed, monosyllabic suffixed, disyllabic suffixed) were tested as within-subject factors, subject were tested as between-subject factors.

Although there was a significant main effect of *tone* type ($F(1,54)=174$, $p <.001$), there was a great between-subject variation and no clear trend was found across subjects.

The main effect of *suffix* type was also found ($F(1,54)=662$, $p <.001$), indicating that the mean peak F0 is significantly higher for suffixed words than for unsuffixed words.

Therefore, we've learn from the data of peak F0 values that the peak is higher in suffixed words than in unsuffixed words both for M-class and R-class. However, no consistent difference was found between M-class and R-class.

The pattern that the peak for M is higher in suffixed words than in unsuffixed words can be interpreted the peak is more raised when followed by a suffix that is low pitch. On the other hand, the pattern that the peak in R is lower for unsuffixed words than for suffixed words can be interpreted that the peak is slightly low due to the influence of the initial low tone target within a syllable. However, the peak can be fully raised in suffixed words, probably because there is enough room for the peak to be fully realized in suffixed words.

### 2.1.3 Discussion

The results of the various analyses of the F0 contours show that the tone class we call M-class and R-class are different in the F0 values and the timing of F0 features. The detailed differences of M and R both in unsuffixed and suffixed words can be expressed with the surface representation in Fig. 2.12. The suffix (*i* 'nominative', *imən* 'if') is underlined and italicized.

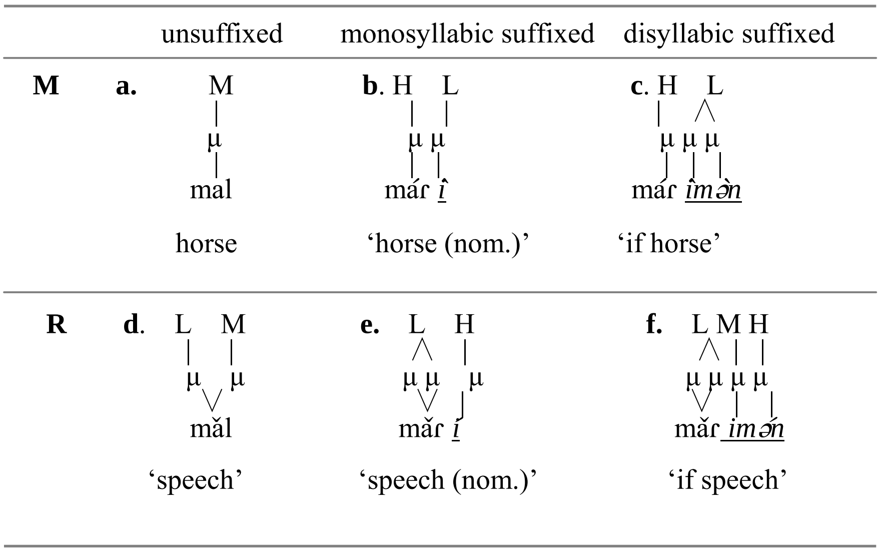



Figure 2.12: Representations of M and R in three suffix type

The empirical support for those representations can be summarized with the following five findings. First, the first syllable of R-tone class has longer syllable durations than that of M-tone class does across suffix type. This can be represented as two moras associated with the initial syllable for R-class.

Second, the fact that the first syllable of R-class has constantly lower initial F0 values than that of M-class does across suffix types can be represented as a low tone associated with the initial mora for R-class.

Third, the finding that F0 peak delay was greater for R-class than for M-class in unsuffixed words can be expressed that the peak is associated with the second mora for R-class.

Fourth, in M-class, the F0 peak values were higher in suffixed words than in unsuffixed words and the peak occurred in the initial syllable. This can be expressed that the peak is a mid tone in unsuffixed words and is a high tone in suffixed words, and a high tone is associated with the first syllable in suffixed words.

Fifth, in R-class, the F0 peak values were higher in suffixed words than in unsuffixed words, and this can be represented that the peak is a mid tone in unsuffixed words and is a high tone in suffixed words. Further, the peak sometimes occurs in the final portion of the second syllable and was held to the early portion of the third syllable but sometimes occurs only in the early portion of the third syllable in disyllabic suffixed R-class. Therefore, I suggest that the pattern of the peak associated with two syllables is not an obligatory one. Instead, it was evident that the peak delay is greater in disyllabic suffixed R-class than in monosyllabic suffixed R-class. Therefore, I propose that the high tone is associated with the second syllable in monosyllabic suffixed R-class but is associated with the third syllable in disyllabic suffixed R-class, as seen in Fig. 2.11. The F0 rising contours of intermediate moras between low toned and high toned ones in disyllabic suffixed R-class are assumed to be a mid tone although their F0 values were not measured in this work.

Consequently, the data in this experiment support the previous transcriptions that the tone class we call R has a lower initial pitch than the class we call M (Choi 1929, Cho 1996) and longer syllable durations (Cho 1996, Kim 1996). In addition, the data provide the acoustic evidence for the tone alternation patterns that the high tone occurs in the root for M-class across the suffix type and that the high tone occurs in the suffix for R in monosyllabic suffixed words (Kim 1996).

However, the findings in this experiment do not support the transcriptions that two tone classes we call R and M are same high tones (Kim 1996), because it was seen that R-class has significantly lower initial F0 values than does M-class across suffix types. The claim that the class we call R is a level low tone (Huh 1955) was not plausible either. While the initial F0 was low for R-class, it has a rising F0 contours and F0 peak in some later portion of the syllable.

Moreover, the pattern in which a high tone is associated with two syllables for disyllabic suffixed R-class (Kim 1996) was, at times, observed for some speakers, but it was not obligatory or regular, even within one speaker. Rather, the peak is realized near the onset of the third syllable, with gradual F0 rising through the second syllable. Therefore, both second syllable and third syllable include regions of high F0, and thus the peak is sometimes stretched over two syllables (final portion of the second syllable and the onset of the third syllable) and sometimes occurs in the onset of the third syllable. Whether the peak occurs in one syllable or is extended over two syllables, the noteworthy pattern is that the R-class has a long peak delay both in unsuffixed words and suffixed words. I understand this long peak delay for the R-class might be induced by the phonetic implementation of a rising tone target. Namely, the peak is in the final portion of a syllable in R-class, and it may take more time for the peak to be fully realized, compared to the simple tone target. This tendency seems to be better displayed in suffixed words, because the tone of the suffix can be weak compared to the tone of the root and thus this context can provide enough room for the peak to be fully implemented. The long peak delay for a rising tone, followed by a weak tone, was also reported in Mandarin (Chen & Xu 2006). When the falling-rising tone is followed by a series of neutral tones (whose tones are weak and dependent on the preceding tone), the F0 peak is realized in the second syllable, following from the falling-rising toned syllable. This pattern is referred to as a "Post-L F0 Raising" in Chen & Xu 2006. Similarly, the peak in a rising tone can be undershot in a syllable but it can be fully realized in a suffixed words, and it might yield a higher peak F0 values for suffixed words than for unsuffixed words in R-class.

## 2.2. EXPERIMENT 1B: M AND R CONTRAST IN PERCEPTION

The production experiment in previous section has demonstrated that M and R are distinct in terms of F0 peak delay, initial F0, and syllable duration. Taking these results into consideration, we can tentatively hypothesize that these three factors play part in contributing to the perception of two tones -M and R-, and no single variable absolutely serves to distinguish M and R. However, the perception study of tonal contrast in this language has not been examined in previous studies.

The purpose of this perception study, therefore, is to determine whether listeners use those acoustic differences in making word identification decision. The experiment tests the perception of isolated synthetic stimuli in which the F0 peak delay, initial F0, and syllable duration have been systematically manipulated relying on the production study.

The production study in previous section shows that F0 peak is later in the syllable, initial F0 is lower, and syllable duration is longer for R than for M. It was seen that the average relative peak delay for 6 speakers was 0.58 for M, suggesting that the peak came at 58% of the syllable duration for M. It was 0.84 for R, indicating that the peak came at 84% of the syllable duration (*see* Table 2.2 in section 2.1). The average syllable duration for six speakers was 215 (ms) and 295 (ms) for M and R, respectively. The average initial F0 value for female was 243 (Hz) and 165 (Hz) for M and R, and the value for male was 121(Hz) and 102 (Hz) for M and R. In principle, all three factors could be strong cues for the distinction between M and R.

Therefore, I hypothesize the late peak delay, low initial F0, and long syllable duration would trigger R response. In sum, the proposed hypotheses can be summarized, as in (16).

(16) Hypothesis

a. The stimuli tend to be identified as the R item if the F0 peak is later in the syllable, initial F0 is lower, and duration is longer.

b. The stimuli tend to be identified as the M item if the F0 peak is earlier, initial F0 is higher, and duration is shorter.

### 2.2.1 Method

#### *2.2.1.1 Stimuli*

To test the hypothesis, two words that differ minimally in lexical tone, /kan/ (M 'taste' - R 'liver'), were chosen for synthesis. They are a minimal pair in which both members are common words in the language. One female native speaker of South Kyungsang Korean produced the isolated M /R tone minimal pairs: /kan/ (M) 'taste', /kan/ (R) 'liver'. Recordings were made in a sound-treated booth in the Phonetics Lab of the Linguistics Department, University of Texas at Austin, using Praat.

With the original sound object, I created a /kan/ continuum using Praat on a PC, by removing the pitch points except for the crucial three points, i.e., initial, peak, and final pitch point. Since the F0 contour in manipulation window of Praat consists of a number of pitch points as shown in Fig. 2.13 below, the pitch contour is needed to be stylized only with the relevant three pitch points.

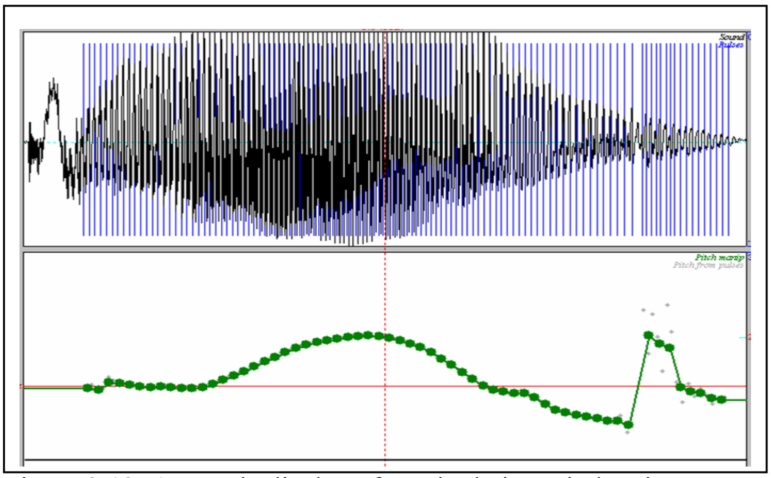


Figure 2.13: A sample display of manipulation window in Praat

The duration and F0 values at the initial, peak, and final points for each tone type in the original recording produced by a female speaker were as follows.

| | Initial point | Peak point | Final point |
|---|---|---|---|
| M | 0 (ms) - 216(Hz) | 201.5 (ms) - 240 (Hz) | 450 (ms) - 136 (Hz) |
| R | 0 (ms) - 170(Hz) | 483.6 (ms) - 224 (Hz) | 620 (ms) - 126 (Hz) |

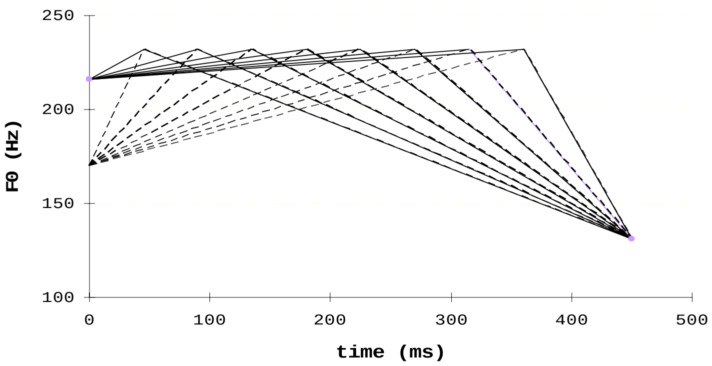
Table 2.10: The F0 values (Hz) and duration (ms) at three points in original recordings

Based on these values in original recording, the stimuli for the perception tests were designed to vary the initial F0 in two steps, i.e., low (170 Hz) and high (216 Hz), and syllable duration in two steps, i.e., short (450 ms) and long (620 ms). The F0 peak delay was varied from 10% to 80% of syllable duration in 10% steps, for a total of 8 variants. The actual values for the F0 peak delay were 45ms, 90ms, 135ms, 180ms, 225ms, 270ms, 315ms, and 360ms for the short syllable duration (450 ms). I manipulated each pitch point according to the proposed values above, by dragging the pitch point, and generated 16 syllables for the short syllable duration. The schematized picture for the manipulated parameters is given in Fig. 2.14 (a) .

The actual F0 peak delay for long syllable (620ms) were 62ms, 124ms, 186ms, 248ms, 310ms, 372ms, 434ms, and 496ms, and its /kan/ continuum also consists of 16 syllables. The schematized picture for the manipulated parameters is given in Fig. 2.14 (b).

___ high initial F0 ---- low initial F0

(a) short syllable




(b) long syllable

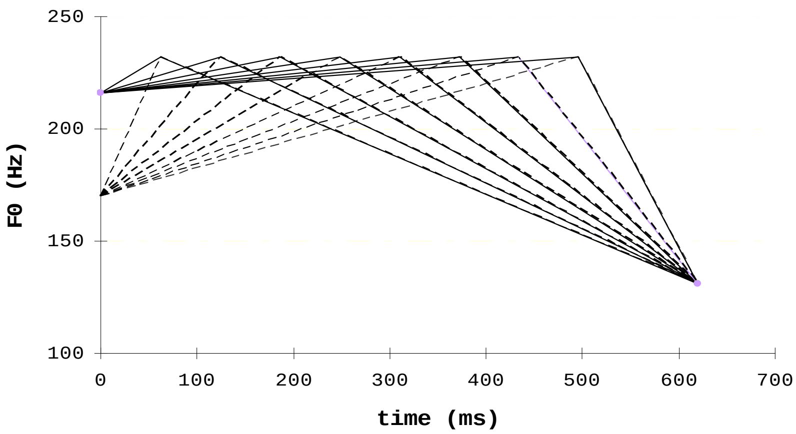



Figure 2.14: A schematic depiction for 36 /kan/ continua

The total /kan/ continuum thus consists of 32 syllables (2 initial F0 * 2 duration * 8 turning point). The peak F0 and final F0 in each stimulus were fixed at constant values, i.e., 232 (Hz) and 131 (Hz) for each syllable, from the average values of M and R in the original recording.

#### *2.2.1.2 Subjects*

Ten adult subjects (aged 20 ~ 58 years old, 5 females and 5 males) from the South Kyungsang area participated. All subjects are linguistically naïve, and all reported that they had no history of hearing impediment. Among the subjects, two participated in preceding production experiment as well. They were born and raised in Pusan, and five of them had lived only in Pusan before coming to the U.S., and five of them had lived in Pusan over 20 years and lived in Seoul for 7 to 10 years.

#### *2.2.1.3 Procedures*

The stimuli were played in randomized order on an audio system using an ALVIN software program on a PC. The detailed information about Alvin program was found at http://homepages.wmich.edu/~hillenbr/. Subjects heard 10 repetitions of each stimulus, a total of 320 tokens. There was a two-second interval between stimuli.

The subjects were instructed to respond to each item as quickly as possible by pressing the button corresponding to pictures and English words for “taste” (M) or “liver” (R) because Korean orthography is exactly same for the two /kan/s. Since encoding picture was not available in ALVIN program to my knowledge, I made the appropriate pictures as seen in Fig. 2.15, and each picture was attached over the corresponding English word on PC screen. I used the Korean soy sauce image for “taste” (M) because we have a common expression “taste with soy sauce”, and liver image for “liver” (R). The pictures used in experiment are given in Fig 2.15.

taste (M)

liver (R)

Figure 2.15: Pictures for response button used in perception experiment of M and R

To avoid misidentification, subjects were asked to pronounce the button labels prior to the experiment. A practice session consisting of 10 test items in each block preceded the test. These practice items provided listeners with end points for each parameter manipulated, as well as stimuli in between. Subject responses were collected by computer using an ALVIN program, and responses for each stimulus were added across speakers.

**2.2.2 Results**

The overall results indicate that R response was triggered when the timing of F0 peak was late, initial F0 was low, and syllable duration was long. The effect of three factors for all ten subjects is illustrated in one graph as in Fig 2.16. It gives the average number of “R responses” out of 10 for all stimuli, arranged according to the “F0 peak delay” and the combination of “initial F0 and syllable duration”, showing all interactions of three factors. The F0 peak delay is represented along the horizontal axis as the relative peak delay (percent), and the average number of R responses is on the vertical axis. The maximum of average number R responses is "10", because one stimulus was heard 10 times to the listeners. The line marked with *empty circle* indicates the "low initial F0-short duration", one with *dark circle* indicates the "low initial F0-long duration", one with *empty square* the "high initial F0-short duration", and one with *dark square* the "high initial F0-long duration".

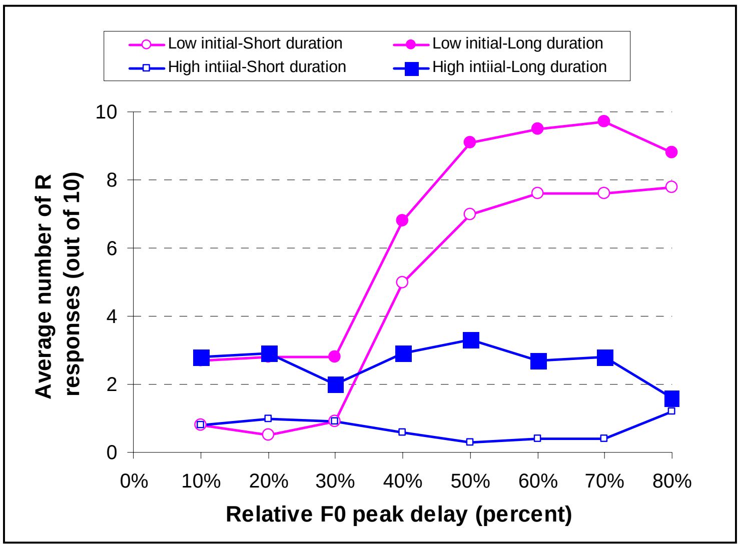



Figure 2.16: The average of R responses according to F0 peak delay, initial F0, and syllable duration (subjects pooled)

To take one example, when the relative peak delay is 70%, initial F0 is low, and syllable duration is long, which is marked with "dark circled line" at 70% timing, the average R responses for that stimulus is almost 10. This indicates that the stimulus was judged as an R item by the subjects for every 10 trial. On the other hand, when the relative peak delay is 10%, initial F0 is high, and syllable duration is short, which is marked with empty squared line at 10% timing, the average R responses is below 2, indicating that this stimulus was judged as an R item less than 2 times for all 10 trials.

Specifically, one notable pattern is that in the dark circled line, the rise begins at 30%, reaching peak slope at 40% with sharp rising curve, and this line has the highest number of R responses. The empty circled line also shows similar pattern with the dark circled line, except that the average number of R responses are slightly lower than that.

This suggests that F0 peak delay have crucial effect on the combination of low initial F0, regardless of duration, triggering more R responses at the later peak delay. However, the F0 peak delay seems to have little effect on the high initial F0. It appears that if initial F0 is high, listeners don't pay attention to timing of peak in identifying M vs R.

It might be because the stimuli with 'high initial F0 and late peak delay' were heard unnatural. Namely, the peak F0 was fixed at 232(Hz) in all stimuli, and the high initial F0 value was 216(Hz). So, the initial F0 value was not that different from the peak F0 value in "high initial F0 stimuli", and this might be heard as a sound with a long peak plateau with a high initial F0, which is not a real word in South Kyungsang Korean. For these stimuli, most subjects seem to respond "M" toned word.

In general, the combination of late peak delay, low initial F0, and long duration triggered highest R responses, and the combination of high initial F0 and short duration has the lowest R responses.

To determine whether these three factors -the timing of the F0 peak, the initial F0, and the duration- significantly affect identification for M and R contrast, logistic regression analysis was conducted. This is appropriate because the dependent variable is binary, M or R. As such, the dependent variable was the dichotomous choice, M or R. M was coded as 0 and R as 1. The independent variables were the peak delay (in milliseconds), initial F0, and duration. The *peak delay* (a) was used instead of percentage of eight variants because continuous variables can provide more informative interpretation in logistic regression. *Initial F0* (b) was coded as 0 for low and as 1 for high. *Syllable duration* (c) was coded as 0 for short and as 1 for long.

The equation for such model is presented in Table 2.2, together with the coefficients (and odds ratios) and p-values. The positive coefficient indicates a direct relationship to the dependent variable. For example, the fact that the coefficients for (a) peak delay are positive indicates that as the peak delay is longer, there is a *more* likelihood for the R response. On the other hand, negative coefficient indicates an inverse relationship with the dependent variable, indicating that as the (b) initial F0 is higher, there is a *less* likelihood for the R response.

| | (a) F0 peak delay | (b) initial F0 | (c) duration | (d) constant |
|---|---|---|---|---|
| Coefficients | **.003(1.003)** | **-1.075 (.341)** | 0.011(1.011) | -0.415(.661) |
| *Equation*: Log odds of R response = (a $\times$ peak delay) + (b $\times$ initial F0) + (c $\times$ duration) + d | | | | |

Table 2. 2: Logistic regression analysis of R response in terms of peak delay, initial F0, and syllable duration across subjects

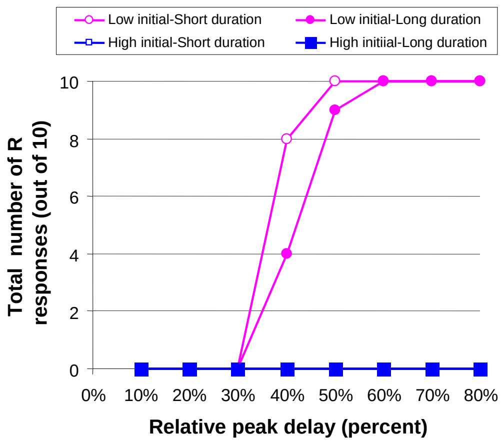
Since this coefficient is in log units, we cannot directly interpret the magnitude of the change. Therefore, the odds ratio converted in decimal units are provided with parenthesis, and it represents the change in the odds of having dependent variable for a one unit of change in the independent variable.

The result, for example, can be interpreted that every one-millisecond-increase in *peak delay*, the odds of R increase by a factor of 1.003 times, i.e., 1 (ms) longer peak delay is 1.003 times *more* likely to be judged as an R. Although 1.003 for peak delay is an extremely small increase, there's a lot of range in milliseconds. For example, if the peak delay is 100 (ms) longer for R than M, then it is 100.3 times more likely to be judged as an R. The stimuli with *high initial F0* is 0.341 times *less* likely to be responded as an R. Likewise, the stimuli with *longer duration* is 1.011 times *more* likely to be responded as an R.

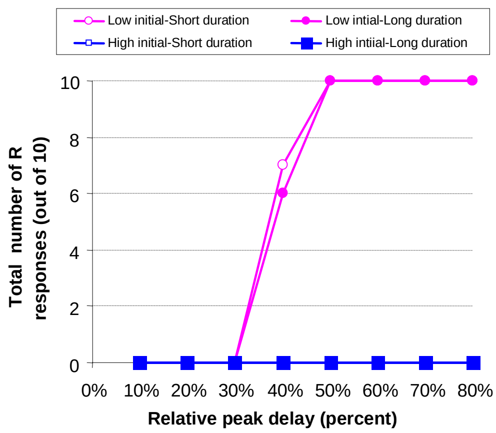
The predictor variables make a significant contribution to the model ($p < .001$) is highlighted in boldface. Only "peak delay" and "initial F0" cues significantly contribute to the perception of M and R contrast.

However, a closer observation of the data for each subject revealed that there are variations among the subjects. Although they tend to show consistent responses to the stimuli with extreme values, they appear to use different strategy for the ambiguous stimulus with intermediate values. Therefore, the effect of three factors is presented by subject in Fig. 2.17 below.

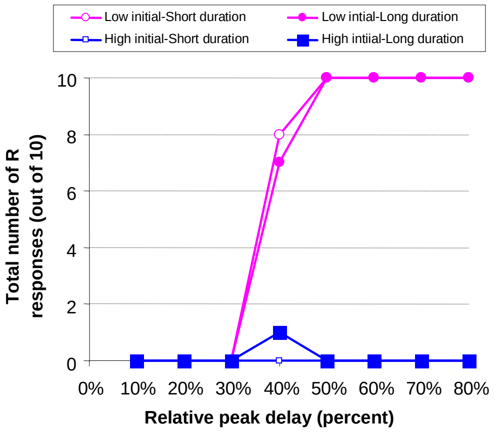
a. subject 1



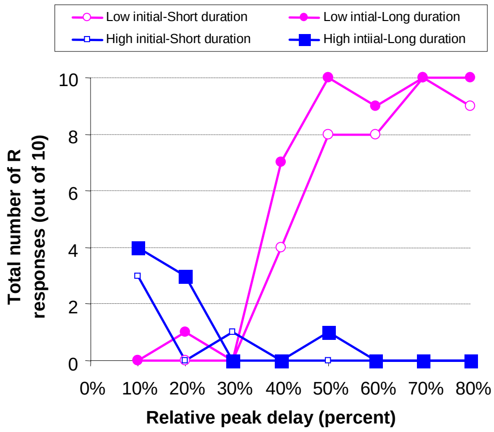

b. subject 2



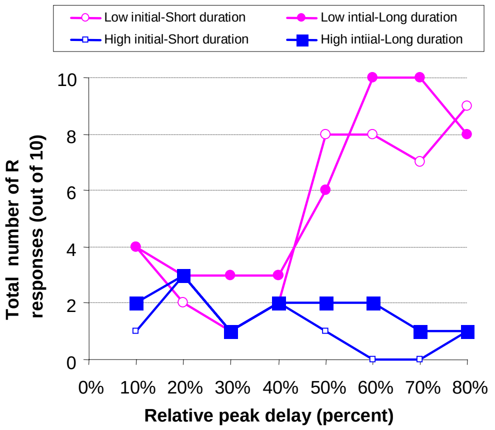
c. subject 3



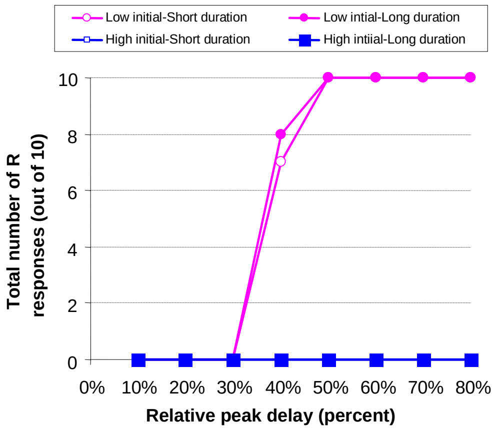

d. subject 4



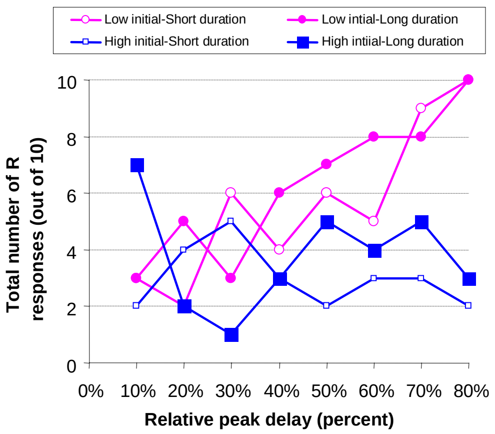

e. subject 5



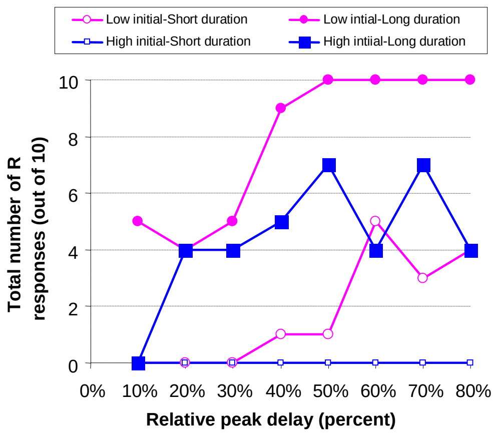

f. subject 6




g. subject 7




h. subject 8



i. subject 9 j. subject 10

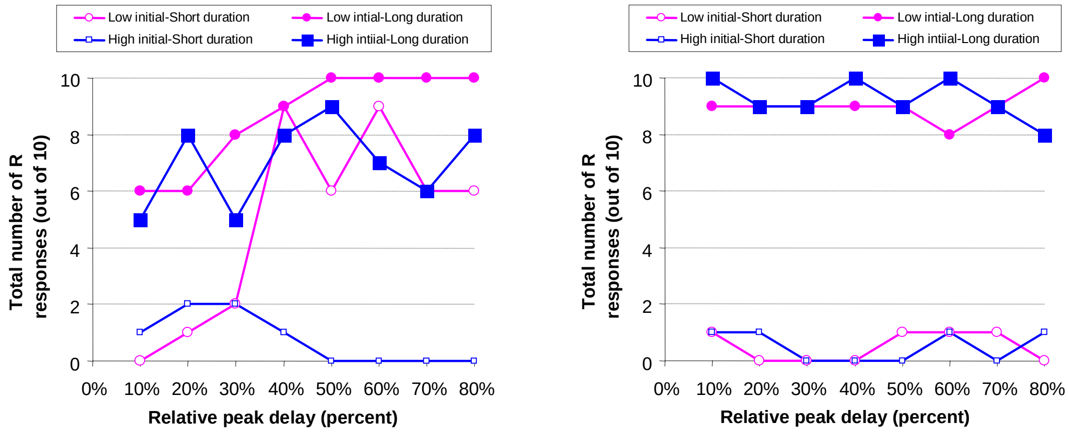



Figure 2.17: The total number of R responses for each subject

Given that the responses of subjects showed great variations, it was possible to group the subjects into three different categories based upon the patterns illustrated in Fig. 2.17 above. When the two lines with a “square” marking a high initial factor are generally constant around “0” value in R responses, and the two lines with a “circle” marking a low initial factor shows a sharp rising, then that subject was grouped in category 1. One subject, who shows an extremely difference pattern that two dark shaped lines have high values and two empty shaped lines have low values, was separately grouped in category 3. The other subjects were grouped in category 2.

We can connect this categorization to the subject’s priority of factors for the ambiguous stimuli. Table 2.3 shows that six subjects (five females and one male) among ten fall into Category 1: subjects weight the F0 cues more heavily than the duration cue. Three subjects (all males) fall into Category 2: three factors were interdependent. One subject (male) falls into Category 3: subjects weight the duration cue more heavily than the F0 cues.

| | Number of subjects | Relative importance |
|---|---|---|
| Category 1 | 6 | timing of F0 peak, initial F0 > duration |
| Category 2 | 3 | timing of F0 peak, initial F0, duration |
| Category 3 | 1 | duration > timing of F0 peak, initial F0 |

Table 2.3: Observed patterns of relative importance for different subjects

In the Category 1 of Fig.2.8. (a-f), one notable pattern is that in the dark circled line, the rise begins at 30%, reaching peak slope at 40% with sharp rising curve, and this line has the highest number of R responses. The empty circled line also shows similar pattern with the dark circled line. This suggests that F0 peak delay have crucial effect on the combination of low initial F0, regardless of duration, triggering more R responses at the later peak delay. However, the F0 peak delay seems to have little effect on the high initial F0. It is notable that subjects in this category do not pay attention to syllable duration as well.

Category 2 in (g-i) shows that R responses were increased as the timing of peak is increased, except for the 'high initial'-'short duration' combination, and that the long duration (dark circle and square) is consistently higher than short duration (empty circle and square) after the 40% of the relative peak delay. The subjects in this category appear that when the stimulus have two factors among three factors in favor of R, then it was judged as an "R", and when the stimulus have two factors in favor of M, then it was judged as an "M". One subject in category 3 seems to consider only duration factor as seen in (j).

To determine whether these three factors -the timing of the F0 peak, the initial F0, and the duration- significantly affect identification for M and R contrast, logistic regression analysis was conducted for each subject. The dependent variable was the dichotomous choice, M or R. M was coded as 0 and R as 1. The independent variables were the peak delay (in milliseconds), initial F0, and duration. The *peak delay* (a) was used instead of percentage of eight variants because continuous variables can provide more informative interpretation in logistic regression. *Initial F0* (b) was coded as 0 for low and as 1 for high. *Syllable duration* (c) was coded as 0 for short and as 1 for long.

The equation for such model is presented in Table 2.4, together with the coefficients (and odds ratios) for each subject. The subjects are arranged in the same order with the Fig. 2.13.

| Subject | (a)<br>F0 peak delay | (b)<br>initial F0 | (c)<br>duration | (d)<br>constant |
|---|---|---|---|---|
| [Category 1] | | | | |
| 1 | **.039(1.04)** | **-11.725(.00)** | -3.579(.03) | -6.503(.00) |
| 2 | **.034(1.03)** | **-10.204(.00)** | -2.646(.07) | -5.661(.00) |
| 3 | **.033(1.03)** | **-10.432(.00)** | -2.357(.09) | -5.404(.01) |
| 4 | **.008(1.01)** | **-3.122(.04)** | -.115(.89) | -1.719(.18) |
| 5 | **.004(1.00)** | **-2.108(.12)** | .132(1.14) | -.803(.45) |
| 6 | **.042(1.04)** | **-13.566(.00)** | -2.447(.09) | -7.206(.00) |
| [Category 2] | | | | |
| 7 | **.003 (1.00)** | **-1.075 (.34)** | .011 (1.01) | -.415(.66) |
| 8 | **.008 (1.01)** | **-2.162 (.12)** | **3.122 (22.69)** | -3.480 (.03) |
| 9 | **.005(1.01)** | **-1.586(.21)** | **2.577(13.16)** | -1.545(.21) |
| [Category 3] | | | | |
| 10 | -.001(.99) | .010(1.01) | **5.153 (172.98)** | -2.508(.08) |

*Equation*: Log odds of R response = (a $\times$ peak delay) + (b $\times$ initial F0) + (c $\times$ duration) + d

Table 2.4: Logistic regression analysis of R response in terms of peak delay, initial F0, and syllable duration

The fact that the coefficients for (a) peak delay are positive for all speakers but for one speaker (subject 10) indicates that as the peak delay is longer, there is a *more* likelihood for the R response. On the other hand, all ten subjects except for one subject (subject 10) showed the negative values for initial F0 (b), indicating that as the (b) initial F0 is higher, there is a *less* likelihood for the R response. The (c) syllable duration has five positive and five negative coefficients.

The odds ratio in log units was converted to decimal units in parenthesis, and it represents the change in the odds of having dependent variable for a one unit of change in the independent variable. The result for speaker 1, to take one example, can be interpreted that every one-millisecond-increase in *peak delay*, the odds of R increase by a factor of 1.04 times, i.e., 1 (ms) longer peak delay is 1.04 times *more* likely to be judged as an R. The stimuli with *high initial F0* is *almost never* likely to be responded as an R. Likewise, the stimuli with *longer duration* is 0.03 times *less* likely to be responded as an R. One pattern to note is that although longer syllable duration was expected to trigger R response more than shorter duration, five of ten subjects showed the opposite direction, suggesting that this factor was not mainly considered to identify M and R for those five subjects, i.e., when two F0 factors are satisfied to be an R response, the duration factor was not counted at all.

The predictor variables make a significant contribution to the model is highlighted in boldface. For example, only F0 cues significantly contribute to the perception of M and R contrast for the subjects in Category 1, and only syllable duration does for the subject in Category 3. Overall, only two subjects in Category 2 show the significant contribution of all three factors to the model. It shows that the different subjects do distinguish the two categories, but apparently on widely differing bases.

### 2.2.3 Discussion

The goal of experiment 1b was to determine whether the acoustic dimensions of peak delay, initial F0, and syllable duration contribute to the perception of South Kyungsang Korean M tone and R tone contrast. The results of the logistic regression analysis seen in Table 2.12 can be interpreted as perceptual models of M and R. The models show that South Kyungsang Korean listeners use these differences when making word identification decisions.

The critical predictor variables were the peak delay and initial F0, i.e., for R to be identified, the perceived target for the peak must be later, and the initial F0 must be lower, than M. The syllable duration also contributes to the model, i.e., for R to be identified, the perceived target for the duration must be longer than M. However, this was a significant factor in the model only for three listeners. Although all three factors were certainly different between the two tones in the acoustic data, the three factors did not equally contribute to the perception of tone contrast for all subjects. Specifically, F0 dimensions were an important cue for 9 of 10 subjects, and syllable duration was an important cue for 3 of 10 subjects. It showed that if one cue is more extreme for a category, other cues can be less so, as observed in perception studies of other languages (Abramson 1975, Gandour 1978, Lin & Repp 1989).

Further, the boundary between the two categories in perception data occurred rather early, i.e., 40% of relative peak delay. The average relative peak delay was 50% for M and 89% for R in the acoustic data (*see* Table 2.2, p.39), namely, a peak F0 located *after* 40% of the syllable duration tends to be judged as an "R" tone. Peaks located *before* the 40% are judged as an "M" tone. The boundary for the response time was consistent with the data of the R response, that is, the response time had a peak at 40% of relative peak delay, suggesting that the stimuli were heard as most ambiguous in this location of the peak.

In addition, M tone perception seems to tolerate more variability overall, while R requires a late peak delay and a low initial F0. This trend also appears to be relevant, as seen in the higher number of M responses than of R responses overall. As noted before, unnatural stimuli were present, including a combination of 'high initial F0 and late peak delay' as well as 'extremely early peak delay'. Most of those stimuli were finally judged as M, although the subjects appeared to have trouble responding. Another possible account is that the M toned words are much more frequent than R in this language (Cho 1996), and it might induce listeners toward "M" response when they heard the unnatural stimulus.

An interesting aspect of the results was the between-subject variation. Subjects distinguished the tone categories, but use different weightings of the cues. We can correlate the variations in the perception task with variations in the production task. Because the subjects participating in the production test were not completely matched with the subjects participating in the perception test, a thorough comparison cannot be made. In the perception test, there were three subjects who showed the significant role of syllable duration. One (Subject 8 in the perception test) of the three also participated in the production test (labeled as Speaker 5 in this test), during which he showed an extremely large difference in syllable duration between M and R. Additionally, the speaker with a markedly lower initial F0 and greater relative peak delay for R in the production test (Speaker 1) showed a stronger cue for the low initial F0 and peak delay factor in the perception test (labeled Subject 3 in this test).

Another between-subject variation is that, although every female subject showed relatively consistent responses, male subjects had large within-subject variations in their responses. One possible interpretation is that the gender of the speaker might influence the perception of the listener. If the speaker and listener are the same gender, the listener may be more familiar with the speaker's voice and thus perceive the stimuli more precisely than would a listener of the other gender.

Overall, the data confirm that South Kyungsang Korean listeners have two tone contrasts, M-class and R-class, which differ in F0 peak delay, initial F0, and syllable duration.

# Chapter 3: Experiment 2

# M and H Contrast

Experiment 2 is concerned with M and H contrast and consists of two production studies (Experiment 2a and 2b) and one perception study (Experiment 2c). Experiment 2a investigates production data for 6 native speakers of South Kyungsang Korean, comparing acoustic measurements of the F0 peak and fall delay, peak plateau, and peak F0 value between M and H both in unsuffixed and suffixed words. M and H with a monosyllabic suffix are compared with monomorphemic HL and HH in Experiment 2b. Experiment 2c tests perception of isolated synthetic stimuli in which acoustic cue was manipulated, based on the findings of production study. 10 native speakers of South Kyungsang Korean are asked to identify the stimuli between monomorphemic HL and HH.

## 3.1 EXPERIMENT 2A: M AND H CONSTRAST IN PRODUCTION

The goal of Experiment 2a is to determine whether there is any acoustic difference between two tone classes (that we call M-class and H-class), and what that difference is.

There have been various descriptions for these two tone classes. In most descriptive studies, the tone class we call M-class has been transcribed as a high tone, and the class we call H as a mid tone (Choi 1929, Huh 1985, Kim-C 1973, Kim-Y 1986). The studies providing phonological analysis suggest that these two tone types are same high tones, distinguished only in suffixed words (Chung 1980, Kim 1996). An instrumental study, Cho (1996), also does not separate these two tone types, describing both of them as a “peak-fall contour”.

However, authors all agree that they have different tone alternation patterns in suffixed words. My observation found that when M is followed by a monosyllabic suffix or a vowel-initial polysyllabic syllabic suffix, the high tone occurs in the root, but when M is followed by a consonant-initial polysyllabic suffix, the high tone occurs in the suffix, as in (17a). In comparison, when H is followed by a suffix, the high tone occurs in the initial syllable of the suffix, regardless of suffix type, as in (17b).

(17) a. <u>M</u>

| | |
|---|---|
| mun | ‘door’ |
| mún-ȉ | ‘door (nom.)’ |
| mún-dò | ‘door also’ |
| mún-əɾò | ‘with the door’ |
| mùn-bə́dà | ‘than door’ |

b. <u>H</u>

| | |
|---|---|
| múl | ‘water’ |
| muɾí | ‘water (nom.) |
| muldó | ‘water also’ |
| muɾésə̀ | ‘in water’ |
| mulbə́dà | ‘than water’ |

However, the pattern shown in (17) is partially contradictory to the transcriptions in earlier literature. For example, Kim (1996) and Ramsey (1975) transcribe H with a suffix as a “doubly-linked H”, such as múɾ<u>í</u> ‘*water (nom.)*’ and múɾ<u>ésə</u> ‘*in water*’. According to Kim (1996), both tone classes are same high tone, but the high tone spreads rightward onto the following syllable in the class we call H, and not in the class we call M. This division within same high tone class is typologically an odd kind of tone contrast. Commonly tone spread is a general phonological pattern, not lexically governed.

Moreover, my pilot study of North Kyungsang Korean (2005) and the preliminary observations on this language cast doubt on the claim that the high tone is associated with two syllables in H-class with a suffix. However, this issue was not thoroughly investigated in those studies.

Therefore, this experiment examines two tone classes we call M-class and H-class both in unsuffixed words and suffixed words, to substantiate the preliminary observation, which is not in accord with earlier studies (Choi 1929, Huh 1955, Kim 1996).

Although my preliminary observations found that it was almost impossible to distinguish these two tone classes from listening to them when produced in isolation, it was also found in pitch tracks that the class we call M has a slightly but consistently lower peak F0 values and longer syllable durations than the class we call H does.

In suffixed words, because the peak comes in the root in case of M-class regardless of the suffix type, we can expect that the peak would come in the initial syllable for the tone class we call M-class across suffix type. By contrast, the peak occurs in the first syllable of the suffix in case of H-class, we can expect that the F0 peak would come in the second syllable for H-class. These all expected differences can be tested on the basis of the F0 peak delays, fall delays, peak F0 values, and syllable durations. For example, if the peak is in the initial syllable for M-class, and in the second syllable for H-class, the peak and fall would be later for H-class than for M-class. The proposed hypotheses for this experiment can be summarized, as in (18), and the difference in the peak location in suffixed words can be schematized, as in Fig. 3.1. The vertical lines in the figure mark the syllable boundary.

(18) Hypotheses

a. F0 *peak value* is higher for tone class we call H-class than for M-class in unsuffixed words.

b. The *test syllable (the first syllable) duration* is longer for M-class than for H-class.

c. In suffixed words, the F0 *peak* and *fall* are attained later in the word in the H-class than in the M-class.

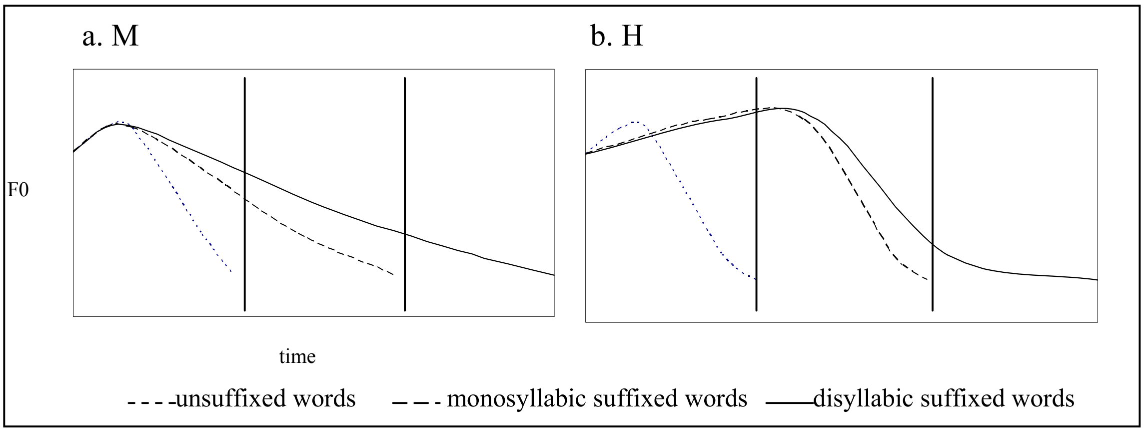



Figure 3.1: The schematic illustrations of F0 contours for M and H in three suffix types

### 3.1.1 Method

#### *3.1.1.1 Materials*

In order to test these hypotheses, the following factors were considered: (i) *tone type*: M and H, (ii) *suffix type:* unsuffixed, monosyllabic suffix, and (vowel-initial) disyllabic suffix. There are 6 conditions (2 tone types * 3 suffix types) and 10 repetitions for each condition, yielding a total of 60 tokens per speaker. Factors which are known to affect the F0 contours and timing, such as segments, phrase-position, adjacent tones, intrinsic vowel height, and syllable structure were controlled as in Experiment 1a (*see* 2.1.1). Due to the difficulty in finding words that meet all these requirements, one word per each condition was chosen. The sentences used as recording material are presented in

Table 3.1. The test word is indicated in bold, word boundary is marked with the space and the morpheme boundary is marked with the '-'.

| M | H |
|---|---|
| a. unsuffixed | a. unsuffixed |
| ʔiɟè **nan** àn-dá | ʔiɟè **nám** àn-dá |
| *now orchid know-dec* | *now other people know-dec* |
| '*Now I know an orchid'* | '*Now I know other people'* |
| b. monosyllabic suffix '*nom.*' | b. monosyllabic suffix |
| ʔiɟè **nán-ì** bò-í-n-dà | ʔiɟè **nam-í** bo̖-i̗-n-dà |
| *now orchid-nom see-pass-pres-dec* | *now other people-nom see-pass-pres-dec* |
| *'Now an orchid is seen'* | *'Now other people is seen'* |
| c. disyllabic suffix '*if~*' | c. disyllabic suffix |
| ʔiɟè **nán-imən** dè-gét-t'à | ʔiɟè **nam-ímən** dè-gét-t'à |
| *now orchid-if fine-fut-dec* | *now other people-if fine-fut-dec* |
| *'Now an orchid would be fine'* | *'Now other people would be fine'* |

Table 3.1: Tone and Suffix types in Phrase-medial used as Recording Material

### *3.1.1.2 Subjects*

Six adult native speakers (aged 21~43, three females and three males) of South Kyungsang Korean (Pusan) produced the materials in this study. The speakers in this experiment are not same with the previous production experiment, Experiment 1a. All speakers are linguistically naïve, and they were born and grew up in Pusan, the capital of South Kyungsang. Three of them lived only in Pusan before coming to the U.S., and three of them had lived in Pusan over 20 years and lived in Seoul for 7 to 10 years.

### *3.1.1.3 Procedures*

The recording procedure was same with the Experiment 1a, and the detailed description can be found in Chapter 2 (*see* 2.1.1.3, p.33).

#### *3.1.1.4 Measurements*

The measurements were also generally same with the Experiments 1a, and the detailed description can be found in Chapter 2 (see 2.1.1.4, p. 33). The one difference is that the initial consonant is in this experiment not utterance initial, and the end of the syllable is marked by the onset of a nasal, not a tap. Since the test syllable is the first syllable of the test word (e.g., the bolded syllable in "**na**n-i *orchid (nom.)*") and begins with an [n], the *onset of the test syllable* was measured at the onset of the marked drop in amplitude in the waveform and a fall to a minimal value of F1 in the spectrogram. For low F0, it was measured at the point at which the local F0 value starts to rise during the test syllable (first syllable).

### 3.1.2 Results

#### *3.1.2.1 The F0 peak value*

To begin with, F0 peak value is tested. The hypothesis was that the F0 peak is higher for the tone class we call H-class than for M-class in unsuffixed words. Figure 3.2 shows the variations in F0 peak value in a box plot graph with tone type and suffix factor compared for each speaker. In unsuffixed words, the median peak value is slightly higher in dark boxes than in empty boxes for four speakers (speaker 1,3,4, and 5), indicating that the peak value is higher in H-class than in M-class for some speakers, but not for some other speakers. Another result is that the peak is generally higher in suffixed words than in unsuffixed words for M-class.

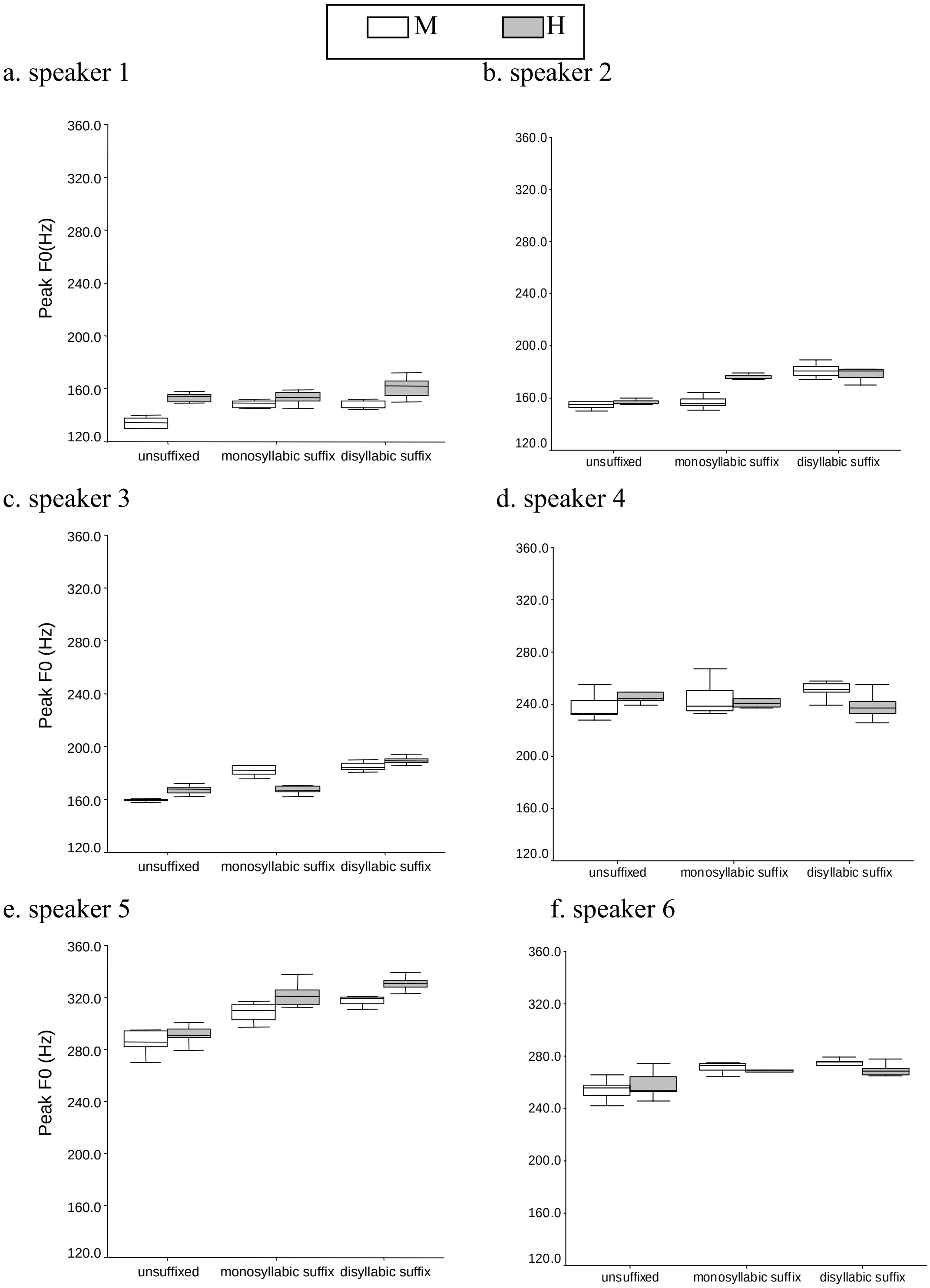



Figure 3.2: *The F0 peak value* for tone type and suffix type for each speaker

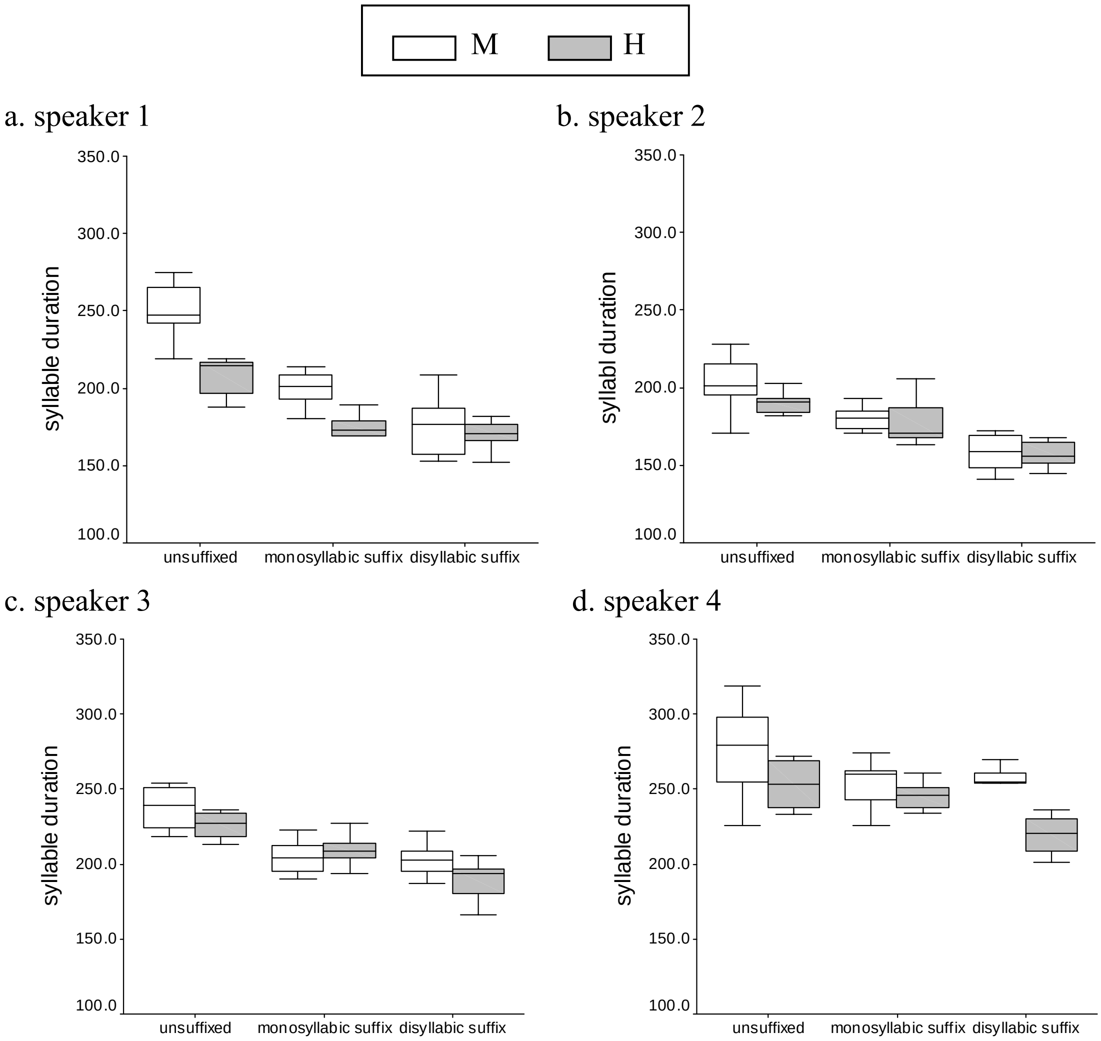
For analysis of peak F0 values, the pitch values in Hz were converted into z-scores separately for each speaker in order to convert the absolute Hz values to locations relative to that subject's pitch range. The repeated measures ANOVA was performed: the dependent variable was the z-scores converted from the peak F0 values in Hz. The tone type (M and H) and suffix type (unsuffixed, monosyllabic suffixed, disyllabic suffixed) were tested as within-subject factors, subject were tested as between-subject factors.

Although there were significant main effects of tone type ($F(1,54)=59$, $p <.001$) and suffix type ($F(1,54)=494$, $p <.001$), these main effects are probably significant only due to the interaction of tone type and suffix type, because H-class is higher than M, not across the board, but only in unsuffixed words. The interaction of the tone type and suffix type was significant ($F(1,54)=26$, $p <.001$).

The Post-hoc tests using Fisher's PLSD showed the following significant differences among six groups: "H, disyllabic suffixed", "H, monosyllabic suffixed", "M, disyllabic suffixed", "M, monosyllabic suffixed" > "H, unsuffixed" > "M, unsuffixed", supporting the hypothesis that the peak F0 value is higher for H-class than for M-class in unsuffixed words. However, there was a variation between subjects. In particular, the significant difference between tone classes in question (M and H in unsuffixed words) was found for four speakers (speaker 1,3,4, and 5), and the mean peak F0 values were lower for H-class in unsuffixed words than for M-class in unsuffixed words for one speaker (speaker 6), and no difference between two tone types for one speaker (speaker 2).

*3.1.2.2 Syllable duration*

The syllable duration is tested in this section. The hypothesis was that the test syllable (the first syllable) duration is longer for M than for H in unsuffixed words. Figure 3.3 shows the variations in test syllable duration in a box plot graph with tone type and suffix factor compared for each speaker. It is seen that the duration is longer for M than for H in unsuffixed words. Also note that the duration is greater for unsuffixed words than for monosyllabic suffixed words, and greater for monosyllabic words than for disyllabic suffixed words.



e. speaker 5 f. speaker 6

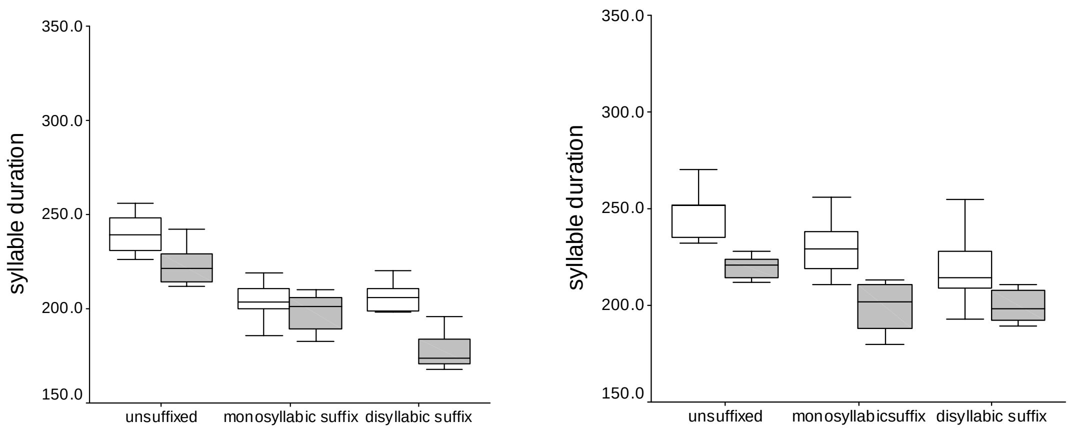



Figure 3.3: *Syllable duration* for tone type and suffix type for each speaker

The data were analyzed statistically by means of repeated measures ANOVA in which syllable durations were dependent variables. Subject was the between-subjects factor, and tone type (M and H) and suffix type (unsuffixed, monosyllabic suffixed, disyllabic suffixed) were the within-subjects factors.

There were a significant main effect of *tone* type (F (1, 54)=18, p <.001), indicating the means are significantly greater for M than for H. However, the effects are not across the board, but are limited to an interaction effect, i.e., unsuffixed M-class. The main effect of *suffix* type (F(1,54)=96, p <.001) was also significant, indicating that the means are significantly greater for unsuffixed words than for suffixed words. The *interaction* of tone type and suffix type was significant (F (1,54)=7, p=.01), with a greater unsuffixed word effect for M. This is probably because the final syllable lengthening effect in unsuffixed words and intrinsic long syllable duration of M are combined for M in unsuffixed words. To see the pair-wise comparison, Post-hoc tests using Fisher's PLSD was conducted. It showed the following significant differences among six groups; "M, unsuffixed" > "H, unsuffixed" > "M, monosyllabic suffixed", "H, monosyllabic suffixed” > “M, disyllabic suffixed”, “H, disyllabic suffixed”. Overall, the hypothesis that the syllable duration is longer for M than in H in unsuffixed words was supported.

### *3.1.2.3 The timing of F0 peak*

In this section, we consider the difference timing of F0 *peak* between the tone type we call M-class and H-class. The F0 peak came later in H than in M in suffixed words. Fig. 3.4 presents typical F0 tracks for the two tone types in unsuffixed words and in monosyllabic suffixed words. The Fig. (3.4 a,b) shows M and Fig. (3.4 c,d) H in suffixed words. The first vertical line marks the onset of the first syllable, the second one marks the offset of the first syllable, and the third one marks the offset of the second syllable. The first arrow indicates the onset of F0 peak, and the second one indicates the offset of F0 peak. Fig. 3.4 shows the F0 peak is in the first syllable for M in suffixed words, as in Fig. 3.4a, while it is in the second syllable for H in suffixed words, as in Fig. 3.4d.

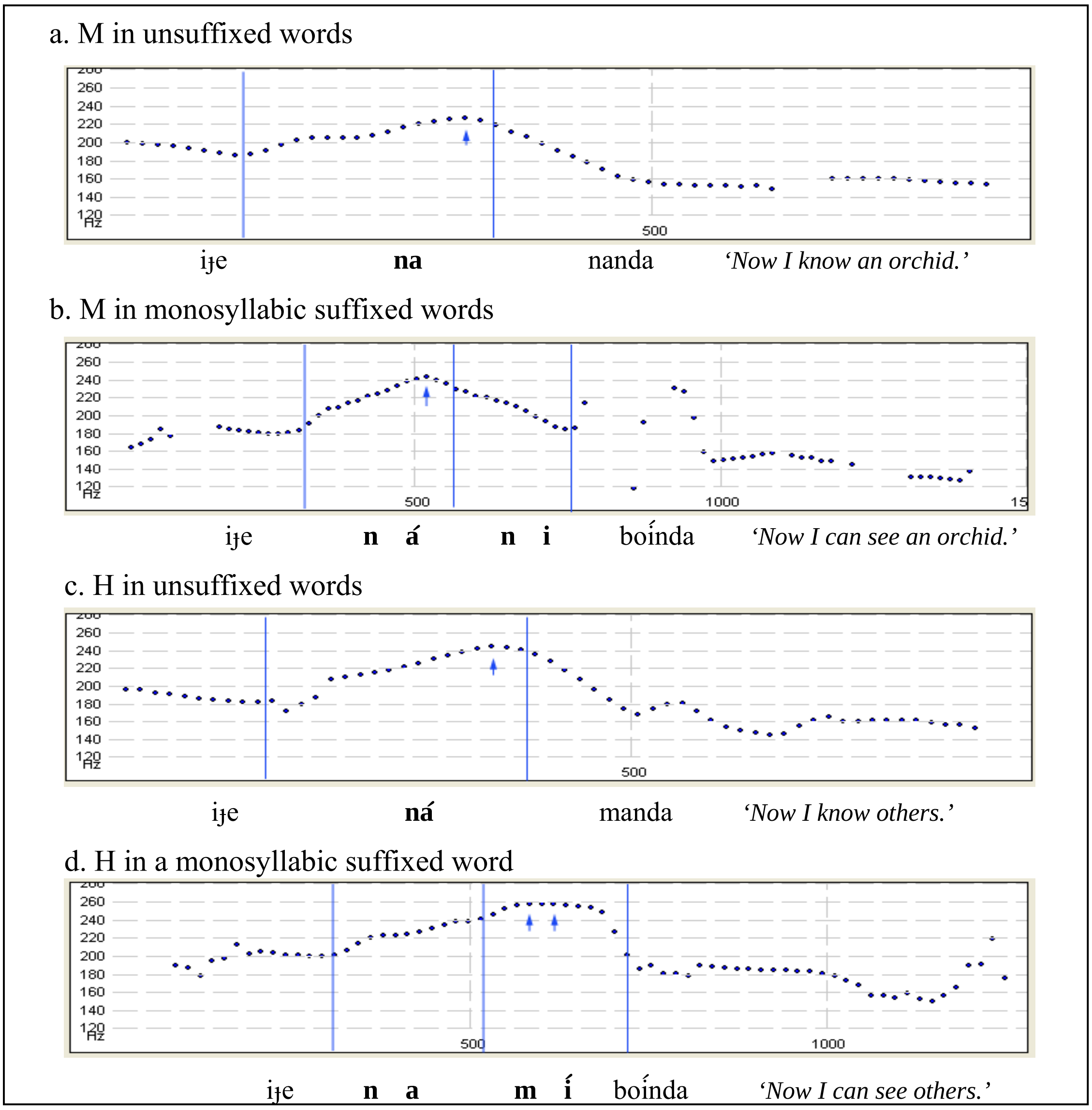
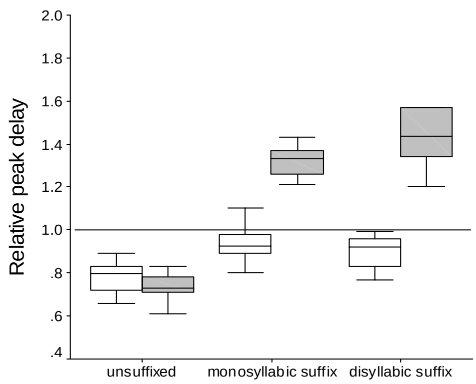


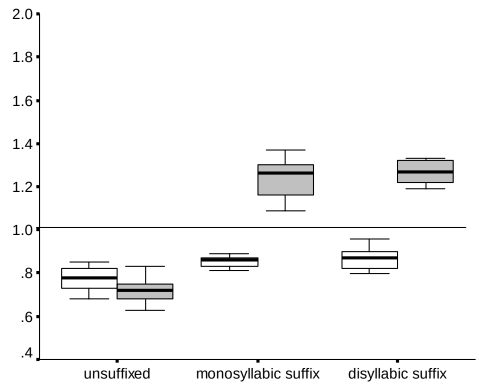
Figure 3.4: Typical effects of M and H in suffixed words on the timing of F0

A quantitative measure of the difference in the timing of F0 peak is the *relative peak delay* (the peak delay divided by the test syllable duration). Fig. 3.5 shows the variation in relative peak delay in a box plot graph with tone type and suffix factor compared for each speaker. The relative peak delay “1” indicates that the peak came precisely *at* the offset of the first syllable, the "less than 1" indicates that the peak came *before* the end of the first syllable, i.e., in the first syllable, and "greater than 1" indicates that the peak came *after* the first syllable. It is seen that the medians in the empty boxes marking M are "less than 1" for all speakers, indicating that the F0 peak came in the first syllable in M across suffix type. The medians in the dark boxes marking H are "greater than 1" for suffixed words, indicating that the F0 peak came after the first syllable for H in suffixed words. It shows that the relative peak delay is greater in H than in M in suffixed words for all speakers.

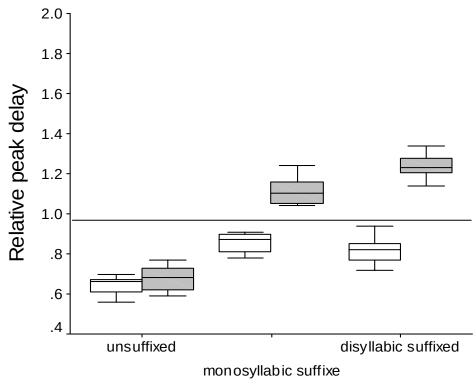

a. speaker 1



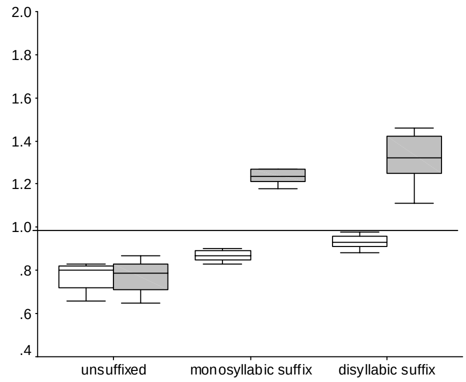

b. speaker 2




c. speaker 3




d. speaker 4



e. speaker 5 f. speaker 6

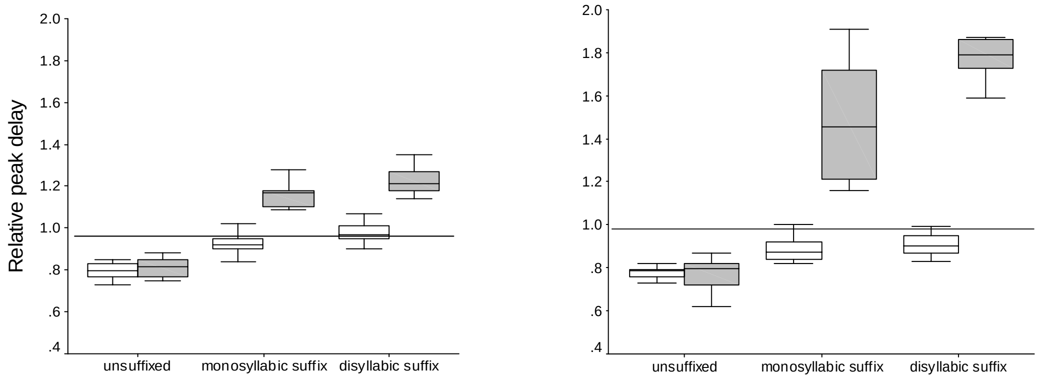



Figure 3.5: *Relative peak delay* for tone type and suffix type for each speaker

The data were analyzed statistically by means of repeated measures ANOVA in which the relative peak delays were dependent measures. Subject was the between-subjects factor, and tone type (M and H) and suffix type (unsuffixed, monosyllabic suffixed, disyllabic suffixed) were the within-subjects factors.

There was a significant main effect of tone type ($F(1, 54)=1060$, $p <.001$), indicating that the means are significantly greater for H than for M. However, the effect is limited in suffixed H-class. The main effect of suffix type was also significant ($F(1,54)=1171$, $p <.001$), indicating that the means are significantly greater for suffixed words than for unsuffixed words. The *interaction* of tone type and suffix type was significant ($F(1,54)=544$, $p=.01$), with a greater suffixed word effect for H. This is because the peak occurs in the second syllable for H in suffixed words.

Post-hoc tests using Fisher’s PLSD showed the following significant differences among six groups; “H, disyllabic suffixed” > “H, monosyllabic suffixed” > “M, disyllabic suffixed”, “M, monosyllabic suffixed” > “M, unsuffixed”, “H, unsuffixed”. The data confirmed that the relative peak delay means are greater for “H, suffixed” than for “M, suffixed”, supporting the hypothesis. However, the significant difference between “H, unsuffixed” and “M, unsuffixed” was not found in the relative peak delay.

### *3.1.2.4 The timing of F0 fall*

We consider the difference of timing of F0 fall between tone classes we call M-class and H-class. A quantitative measure of the difference in the timing of F0 fall is the *relative fall delay* (the fall delay divided by the test syllable duration). The test syllable was the first syllable of the test word. The important result about relative fall delay is that it reflects the pattern of relative peak delay.

Fig. 3.6 shows the variation in relative fall delay in a box plot graph with tone type and suffix type compared for each speaker. It is seen that the medians in the empty boxes marking M are “less than 1” for all speakers but for one speaker (speaker 1), indicating that the F0 fall came in the first syllable in M across suffix type. The medians in the dark boxes marking H are “greater than 1” for suffixed words, indicating that the F0 fall came after the first syllable for H in suffixed words. It shows that the relative fall delay is greater in H than in M in suffixed words for all speakers.

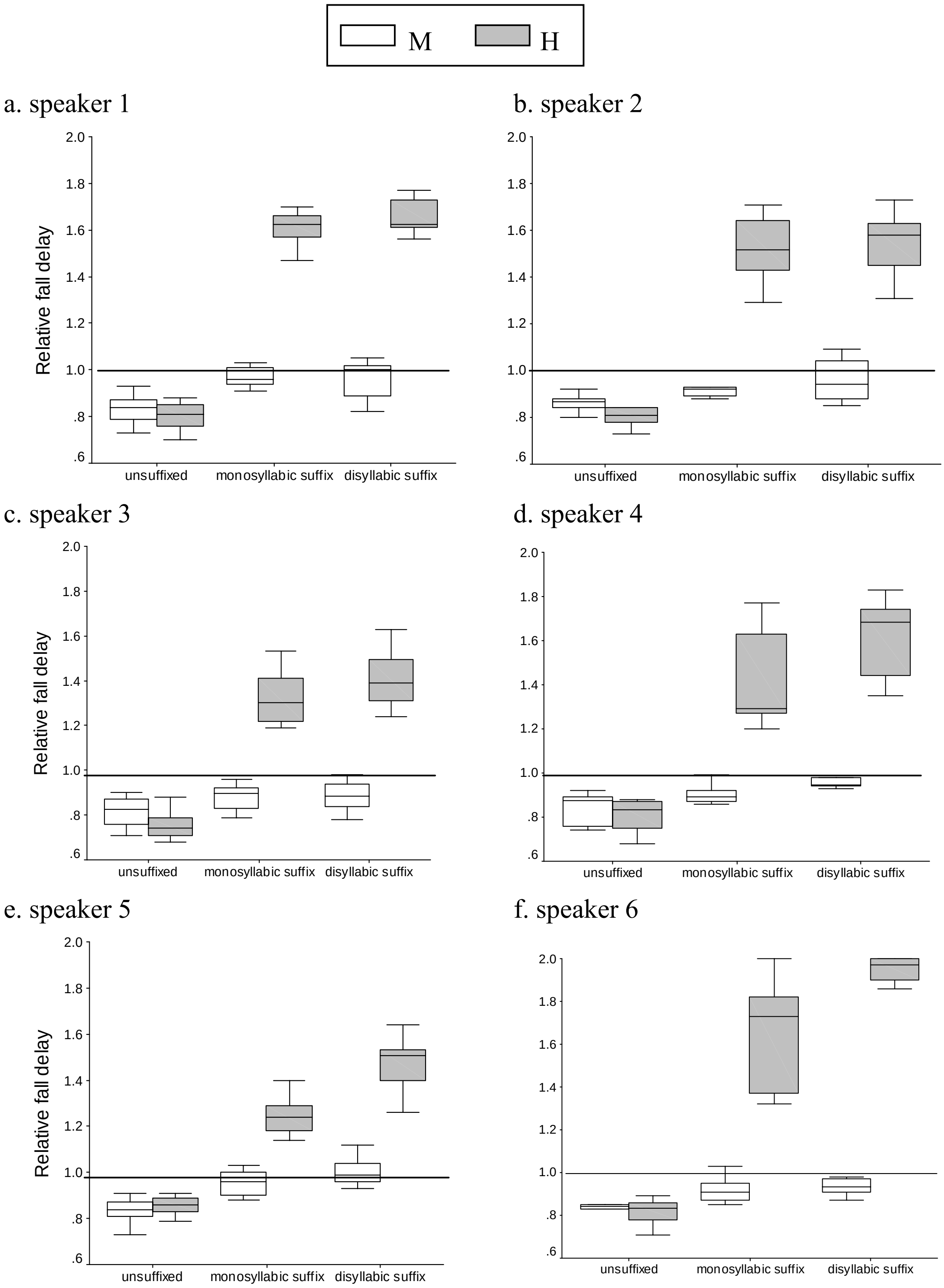



Figure 3.6: *Relative fall delay* by tone type (M and H) and suffix for each speaker

The data were analyzed statistically by means of repeated measures ANOVA in which the relative fall delays were dependent measures. Subject was the between-subjects factor, and tone type (M and H) and suffix type (unsuffixed, monosyllabic suffixed, disyllabic suffixed) were the within-subjects factors.

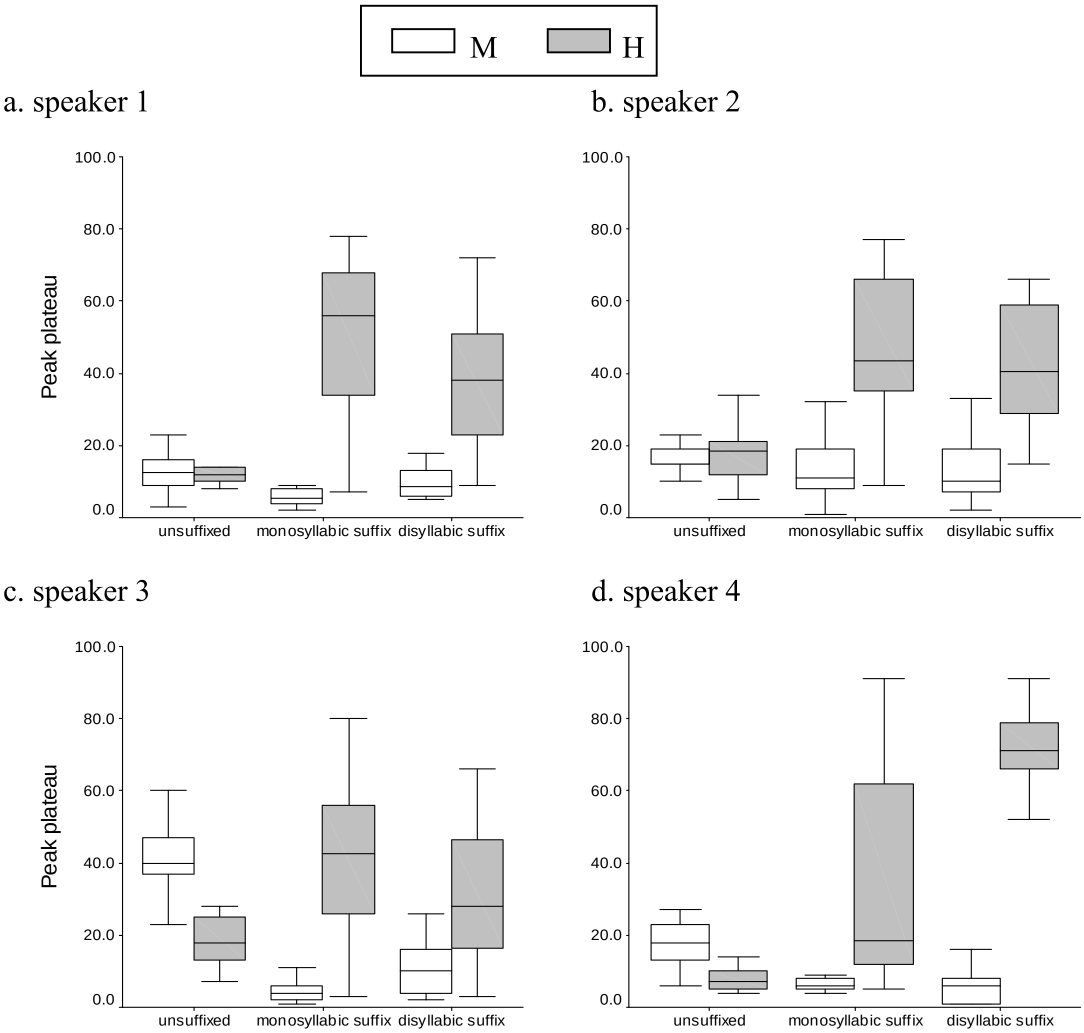
There was a significant main effect of *tone* type, $F (1, 54)=1070$, $p <.001$, indicating that the means are significantly greater for H than for M. However, the effect is limited in suffixed H-class. The main effect of *suffix* type was also significant ($F(1,54)=1563$, $p <.001$), indicating that the means are significantly greater for suffixed words than for unsuffixed words. The *interaction* of tone type and suffix type was significant ($F (1,54)=723$, $p=.01$), with a greater suffixed word effect for H. This is because the high tone occurs in the second syllable for H in suffixed words.

Post-hoc tests using Fisher's PLSD showed the following significant differences among six groups; "H, disyllabic suffixed" > "H, monosyllabic suffixed" > "M, disyllabic suffixed", "M, monosyllabic suffixed" > "M, unsuffixed" > "H, unsuffixed". The data of relative fall delay is generally similar in pattern with the relative peak delay, indicating that the F0 fall came in the initial syllable for M, and in the second syllable for H in suffixed words.

*3.1.2.3 Peak plateau*

The Fig. 3.7 shows the variation in *peak plateau* (the onset of the F0 fall - the onset of the F0 peak) in a box plot graph with tone type and suffix factor for each speaker. It is seen that the peak plateau was greater for H in suffixed words than for others, but the variation range of the peak plateau is considerably large. That is, the difference of the $90^{th}$ percentile and the $10^{th}$ is great, indicating that the peak plateau was sometimes very long and sometimes very short for the H in suffixed words. In addition, it is notable that the peak plateaus are all very short relative to syllable duration.

\



e. speaker 5 f. speaker 6

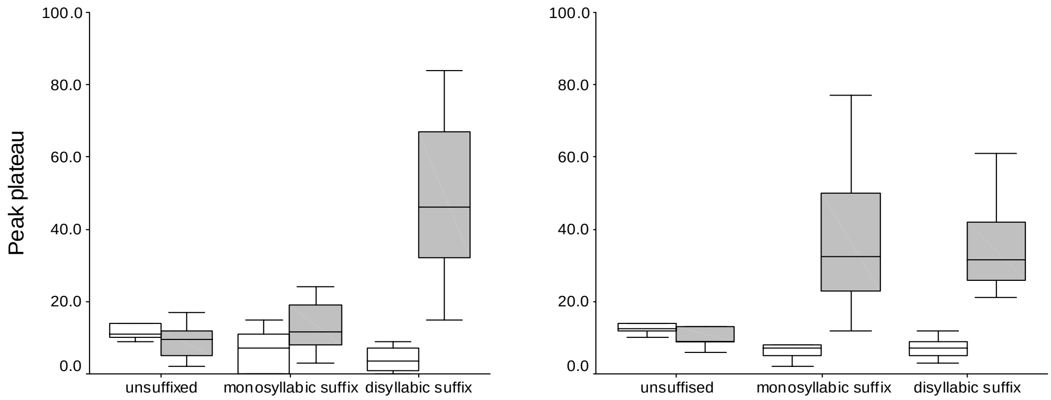



Figure 3.7: *Peak plateau* by tone and suffix for each speaker

The repeated measures ANOVA was used for a statistical test. The peak plateau was the dependent variable. There were two within-subject factors: tone type (M and H) and suffix type (unsuffixed, monosyllabic suffixed, and disyllabic suffixed). Subjects were considered as a between-subject factor.

The result found the significant main effect of *tone* type ($F(1, 54)=113$, $p<.001$), indicating that the mean peak plateau is significantly greater for H-class than for M-class. However, the effect is not overall difference, but limited in suffixed H-class. The main effect of *suffix* type was also significant ($F(1,54)=13$, $p=.001$), indicating that the means are significantly greater for suffixed words than for unsuffixed words. The *interaction* of tone type and suffix type was significant ($F(1,54)=142$, $p<.001$), with a greater suffixed word effect for H. This is because both the first syllable and the second syllable include quite high F0 range for H-class in suffixed words.

A pair-wise comparison by a post hoc test showed that the mean peak plateau was significantly greater for "H, suffixed" than for others. This quite long peak plateau for H in suffixed words can explain why H tone in suffixed words is heard as tone spread.

However, the peak plateau is not necessarily laid upon “two” syllables and the pattern is not obligatory.

### 3.1.2 Discussion

The findings of the present study seem to support the hypotheses that there are acoustic differences between two tone classes we call M-class and H-class, in the F0 peak and fall delays, peak F0 values, and syllable duration. The detailed differences of M and H both in unsuffixed and suffixed words can be expressed with the surface representation in Fig. 2.11. The suffix (<u>*i*</u> ‘nominative’, <u>*imən*</u> ‘if’) is underlined and italicized.

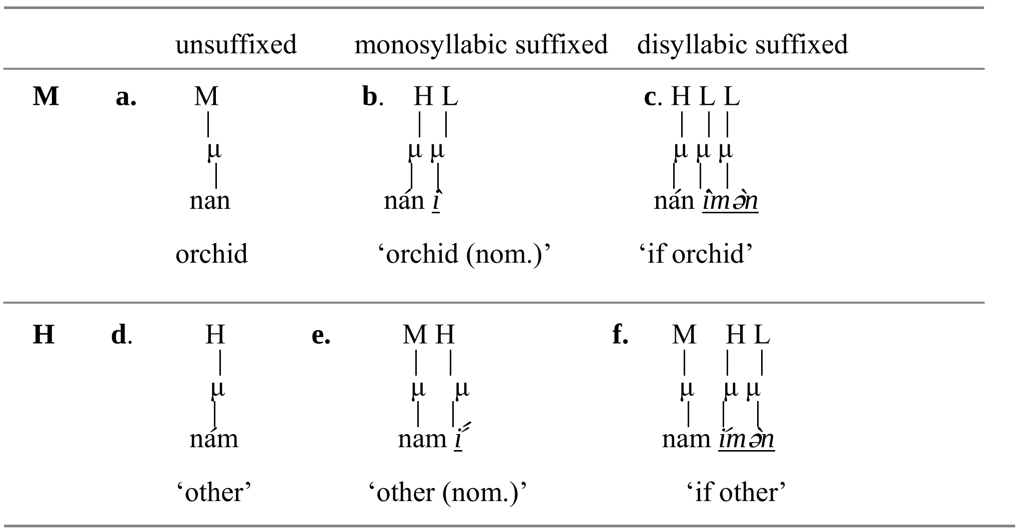


Figure 3.8: Representations of M and H in three suffix types

The empirical support for those representations can be summarized with two main findings. First, in unsuffixed words, the peak F0 values were higher for H than for M, but the statistical significance was found only for four speakers of six speakers, and the difference was not great. Similarly, the syllable duration was significantly longer for M-class than for H-class for all six speakers, but their difference in syllable duration was just 19(ms) while the difference was 79(ms) between tone classes we call M and R in Experiment 1a.

Therefore, it is suggested in Fig. 3.8 that both tone classes are associated with one mora while the tone class we call R was suggested to be associated with two moras in Experiment 1a. In addition, the slightly lower peak F0 values for the M-class than for the H-class are represented that M-class is associated with the mid tone while H-class is associated with the high tone. This representation also reflects the lower F0 peak values for M in unsuffixed than for M in suffixed words across subjects. However, given that the difference in peak F0 values between mid tone and high tone is quite small, further investigation is needed to verify the distinction between M-class and H-class in unsuffixed words.

Second, in suffixed words, it was found that the peak occurs in the initial syllable of the suffix (the second syllable) for the tone class we call H-class, but in the root (first syllable) for M-class. This data is reflected in Fig. 3.8 that the high tone is associated with the first mora for the M-class, but with the second mora for the H-class. The peak plateau in suffixed words was sometimes longer for H-class because the F0 peak was realized with gradual F0 rising contour through the initial syllable and thus the initial syllable and the second syllable embrace regions of high F0. Therefore, it is suggested that the first mora is associated with the mid tone, which is slightly lower than the high tone.

To summarize, the proposed representation in unsuffixed words is different with either previous claim that the tone class we call M-class is a high tone and the H-class is a mid tone (Choi 1929, Huh 1955) or that the two tones are same high tone (Kim 1996). Since the difference in peak F0 values and syllable durations were found, their differences were illustrated in different tone targets in this work. However, the contrast between these two tone classes in unsuffixed words deserves further investigation because their differences were not remarkable as shown in the contrast between M-class and R-class in Experiment 1a, and the consistent pattern across the speakers were not found.

The representation in suffixed words is also different with the previous claim that the high tone is associated with both first and second syllables for H-class in suffixed words (Kim 1996). Although the pattern, in which the peak was laid upon two syllables, was found, at times, for some speakers, it was not obligatory or regular, even within one speaker. Rather, the peak is realized near the onset of the second syllable, with gradual F0 rising through the initial syllable. Therefore, both the first syllable and second syllable include regions of high F0, and thus the peak is sometimes stretched over two syllables (final portion of the first syllable and the onset of the second syllable) and sometimes occurs in the onset of the second syllable. This pattern is similar with the one shown in R-class in disyllabic suffixed words in Experiment 1a.

Therefore, same explanation can be applied to this pattern. That is, the peak delay for H-class in suffixed words might be due to the phonetic implementation of the high tone target. It would take longer time to reach a high peak (H-class) than a lower peak (M-class), and this might lead to a longer peak delay for H-class than for M-class in suffixed words. This difference might not be remarkable in a syllable, and thus the peak delay might not be different between H-class and M-class in unsuffixed words, but the distinction can be maximized in suffixed words because the peak can be fully implemented in suffixed words. This peak delay pattern for a high tone target, i.e., the peak occurs after the syllable that carries the high tone, has been reported in English, Spanish, Chichewa, Mandarin, and other languages (Silverman & Pierrehumbert 1990, Prieto et al., 1995, Myers 1999, Xu 2001, 2006).

One might wonder whether the phonetic realization of a high tone that is linked two syllables in this language may be the usual case, that is, the peak is realized only in one syllable. This possibility will be further discussed in the next section by exploring how the F0 contours would be realized for a monomorphemic double H tone (i.e., tone class we call HH). The monomorphemic HH and H with a suffix (H + suffix) have been described as the same representation in the literature (Ramsey 1975, Kim 1996). The F0 contours of these two tones will be compared in the next experiment.

## 3.2 EXPERIMENT 2B: MONOMORPHEMIC AND BIMORPHEMIC WORDS IN PRODUCTION

Experiment 2b is designed to compare monomorphemic words and bimorphemic words. One specific comparison is between monomorphemic “HL” and “M with a monosyllabic suffix (M + suffix)”. The other comparison is between monomorphemic “HH” and “H with a monosyllabic suffix (H + suffix)”.

M and H with a monosyllabic suffix have been described in the same way with a monomorphemic HL and a monomorphemic HH in previous studies (Ramsey 1975, Huh 1985, Kim 1996). According to these studies, for example, M with a monosyllabic suffix (M+suffix) is transcribed as H+L (e.g., sɨ́l+ɨ̀ ‘alcohol (nom.)), which is a same representation with monomorphemic HL (e.g., hɨ́nə̀l ‘sky’). This pair is arbitrarily referred to as a “tone type 1” in this work. Similarly, H with a monosyllabic suffix is transcribed as H+H (e.g., mɨ́l+ɨ́ ‘water (nom.)), which is a same representation with monomorphemic HH (e.g., kɨ́ɾə́m ‘cloud’). We will call this pair “tone type 2” for convenience’s sake.

As briefly noted in introduction chapter, however, the F0 peak extended over two syllables is observed for monomorphemic HH, while such a pattern was not regularly observed for “H + suffix”. The peak was generally in the second syllable and thus the peak plateau in “H + suffix” was not as long as in monomorphemic HH.

One goal is, therefore, to verify this preliminary observation in my pilot study. The preliminary observation found that the F0 peak is later in “H + suffix” than in monomorphemic HH, because the peak occurs in the suffix for “H + suffix” while the peak occurs both in the initial syllable and the second syllable. Hence, the peak plateau would be longer in monomorphemic HH than in “H + suffix”.

Another goal is to examine whether there is any acoustic difference between monomorphemic HL and HH, since there has been no previous study to provide the physical evidence for these two tone classes. It was found in the preliminary observation that the difference between HL and HH is that a F0 peak is in the first syllable for HL, and a peak is in the first and the second syllables for HH. This predicts the difference in the timing of F0 fall. The F0 fall would be later in HH than in HL because the peak is extended to the second syllable for HH. Thus the peak plateau is also longer for HH than for HL. In addition, my observations found that the peak is higher for HL than for HH. The proposed hypotheses can be summarized as in (19), and a schematic depiction of those hypotheses is given in Fig. 3.9.

(19) Hypotheses

a. The F0 *peak* is attained later in the word in bimorphemic "H+suffix" than in monomorphemic HH.

b. The F0 *fall* is attained later in the word in HH than in HL.

c. The *peak plateau* is longer in the word in HH than in HL, and longer in HH than in "H+ suffix".

d. The *peak* is higher for HL than for HH.

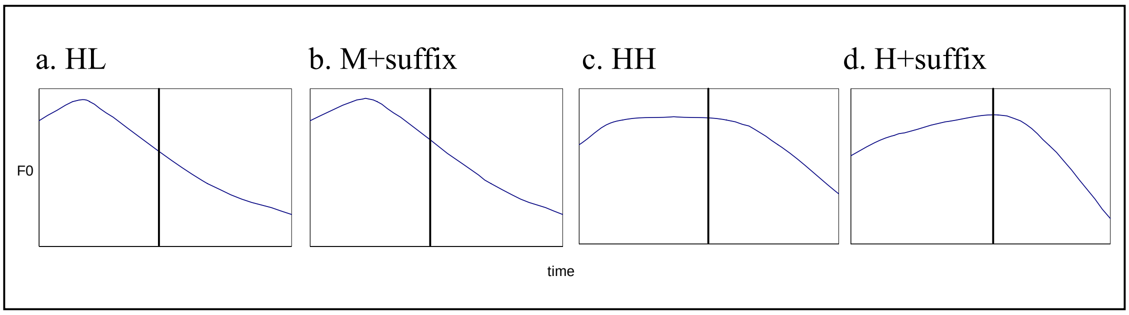



Figure 3.9: The schematic illustrations of F0 contours for HL, M+suffix, HH, and H+suffix

### 3.2.1 Methods

In order to test these hypotheses, two factors, i.e., tone type and morpheme type, were considered. There are two *tone types*, i.e., tone type 1 (HL, M+suffix), which is transcribed as a HL in previous studies (Ramsey 1975, Kim 1996) and tone type 2 (HH, H+suffix), which is transcribed as a HH in previous studies (Ramsey 1975, Kim 1996). There are also two *morpheme types,* i.e., monomorpheme and bimorpheme. Thus, there are 4 conditions (2 tone types * 2 morphemes) and 10 repetitions for each condition, yielding a total of 40 tokens per speaker. Factors which are known to affect the F0 contours and timing, such as segments, phrase-position, adjacent tones, and syllable structure were controlled as in previous experiments (*see* 2.1.1, p.32). Given that HL, HH, "M+ suffix", and "H+ suffix" are to be compared with each other, their intrinsic vowel height should be controlled among them. However, due to the difficulty in finding words whose vowel structures are same across the all four types, the condition was controlled only within the same morpheme type. The test materials used in this experiment are presented in Table 3.2.

| Monomorphemic words | Bimorphemic words |
|---|---|
| a. HL<br>iɟe mɔ́ɾè bò-í-n-dà<br>*now sand see-pass-pres-dec*<br>*'Now sand is seen.'* | b. M+suffix<br>iɟe nán-i bò-í-n-dà<br>*now orchid-nom see-pass-pres-dec*<br>*'Now an orchid is seen'* |
| c. HH<br>iɟe mɔ́ɾé mòí-n-dà<br>*now the day after tomorrow gather - hab-dec*<br>*'We gather the day after tomorrow.'* | d. H+suffix<br>iɟe nam-í bò-í-n-dà<br>*now other people-nom see-pass-pres-dec*<br>*'Now other people is seen'* |

Table 3.2: Test Materials

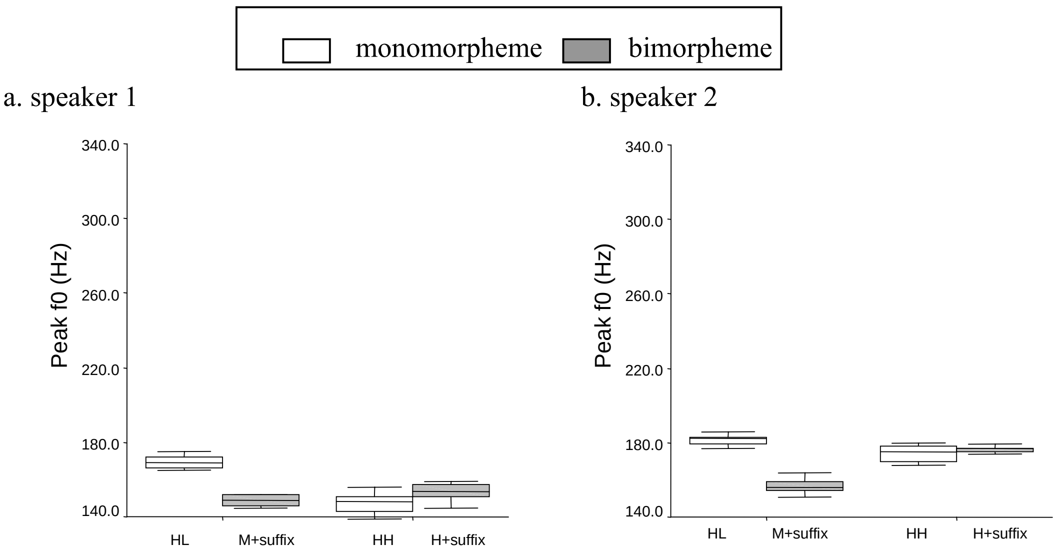
Six adult native of South Kyungsang Korean (Pusan) who participated in Experiment 2a produced the materials in this study as well. The experimental procedure including measurement is same as in previous experiment (*see* 2.1.1, p.33).

### 3.2.2 Results

*3.2.2.1 The F0 peak value*

In this section, the differences in F0 peak values between HL, “M+suffix”, HH, and “H+suffix” are tested. The hypothesis was that the F0 peak is higher in HL than in HH. Fig. 3.10 shows the variations in F0 peak value in a box plot graph with tone type and morpheme type compared for each speaker. F0 peak values (Hz) are indicated in the y-axis, and the morpheme and tone type are arranged in the x-axis.

It is seen that the median lines in monomorpheme are higher in empty boxes than in lined boxes for all speakers but for one (speaker 3), suggesting that F0 peak is higher in HL than in HH. The peak difference between bimorphemic was already discussed in previous section.



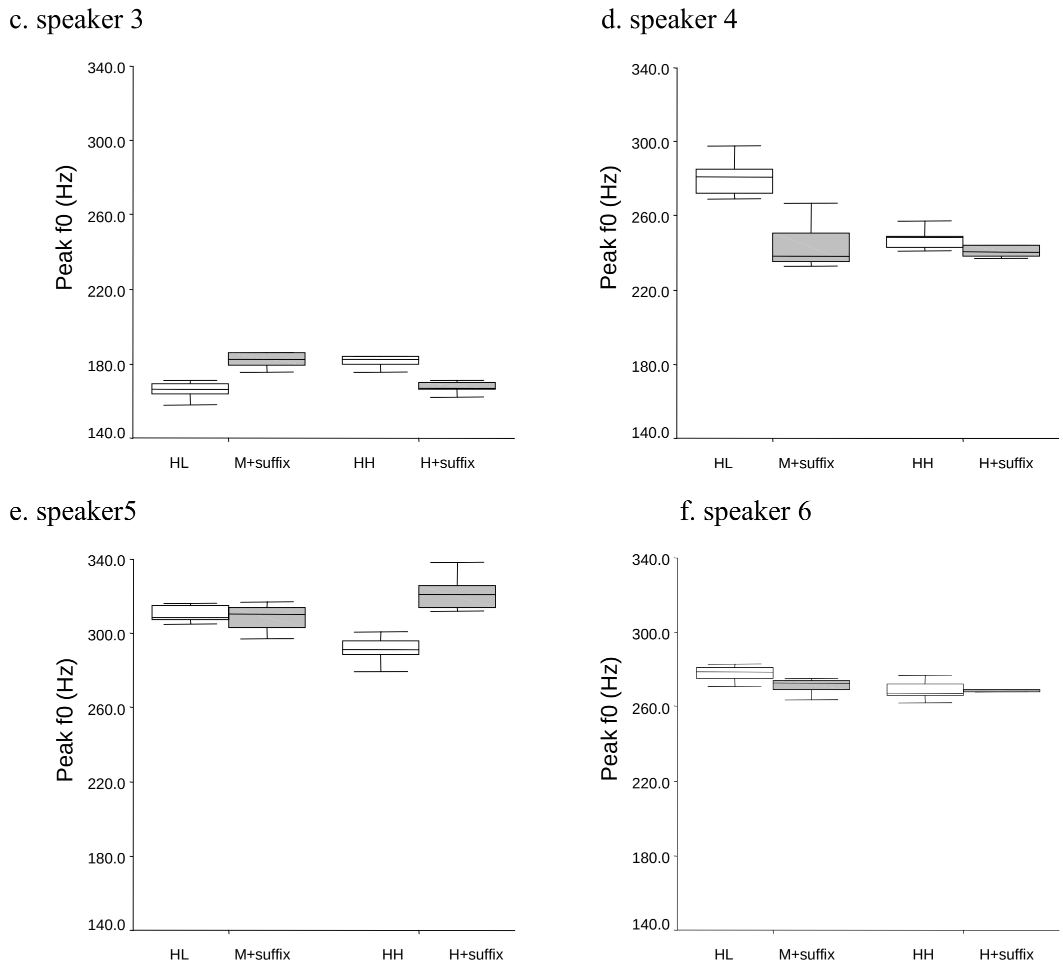



Figure. 3.10. The *F0 peak* for HL, M+suffix, HH, and H+suffix

For analysis of peak F0 values, the pitch values in Hz were converted into z-scores in order to compare across speakers with radically different F0 ranges (e.g., men and women). The repeated measures ANOVA was performed: the dependent variable was the z-scores converted from the peak F0 values in Hz. Tone type (single H and doubly-linked H) and morpheme type (monomorpheme, bimorpheme) were tested as within-subject factors, subject were tested as between-subject factors.

Although there were significant main effects of tone type ($F(1,54)=24$, $p < .001$) and morpheme type ($F(1,54)=25$, $p < .001$), these main effects are probably significant only due to the interaction of tone type and morpheme type, because “tone type 1” is higher than “tone type 2”, not across the board, but only in monomorphemic words, i.e., HL. The interaction of the tone type and morpheme type was significant ($F(1,54)=51$, $p <.001$). A pair-wise comparison by a post hoc test showed that the mean peak F0 values were significantly greater for monomorpheme HL than for the other groups. It supports the hypothesis that the peak is higher for HL than for HH. This can be interpreted that the peak is more raised when followed by a tone that has low F0 value (Gandour et al., 1994, Xu 1997, Laniran & Clements 2003). In addition, the peak is higher for HL than for “M+suffix”, but there’s a variation between subjects. The peak was significantly greater for HL than for “M+suffix” for four speakers (speaker 1,2,4, and 6).

#### *3.2.2.2 The timing of F0 peak*

In this section, the peak delay difference between monomorphemic HH and bimorphemic H+ suffix is explored. The hypothesis was that the peak is later in “H+ suffix” than in HH, and the results confirmed this prediction. Fig.3.11. presents the typical F0 tracks for the four tone types. Fig 3.11 (a) is for the monomorphemic HL, (b) is for HH, (c) is for the bimorphemic “M+suffix”, and (d) is for “H+suffix”.

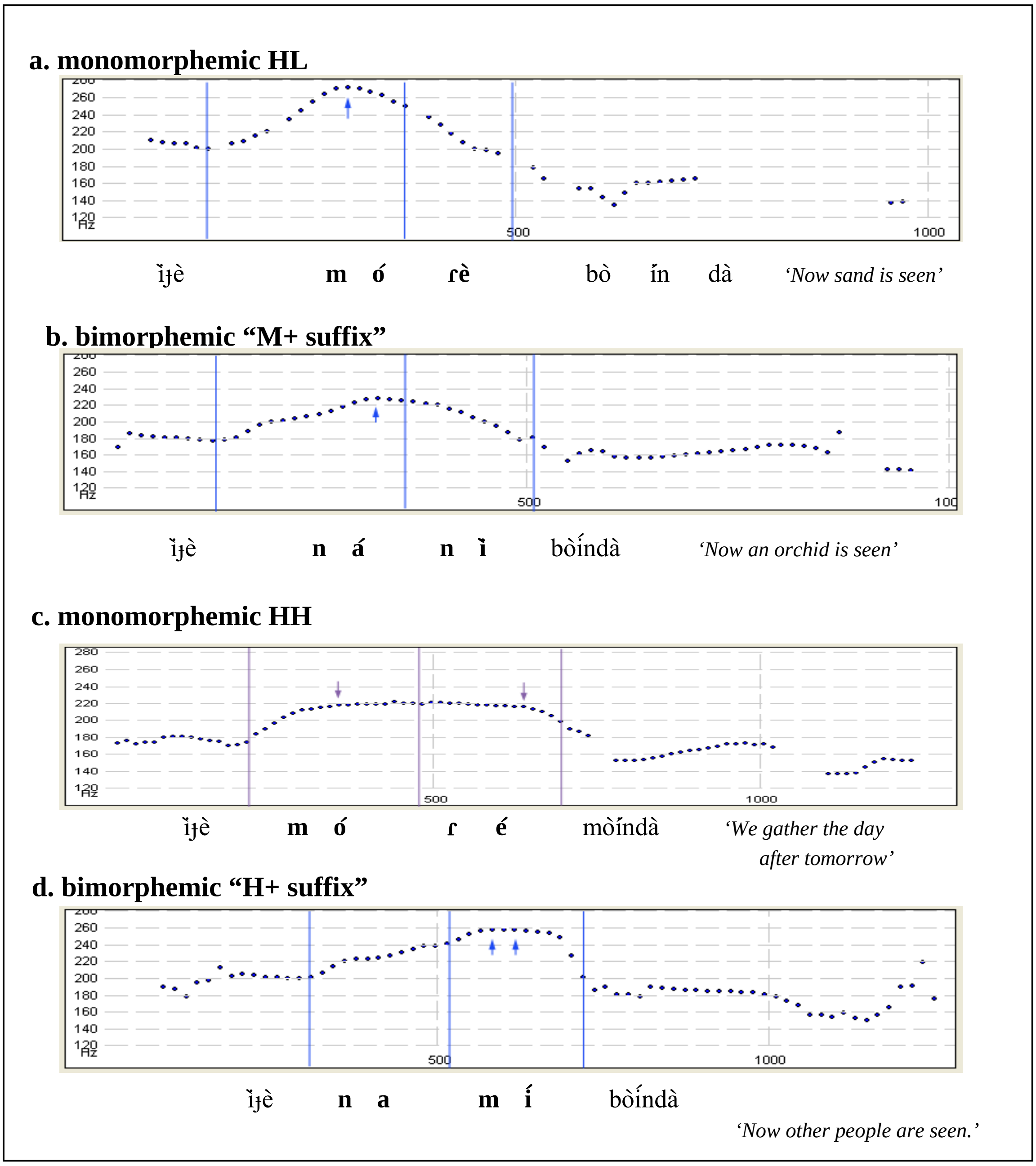
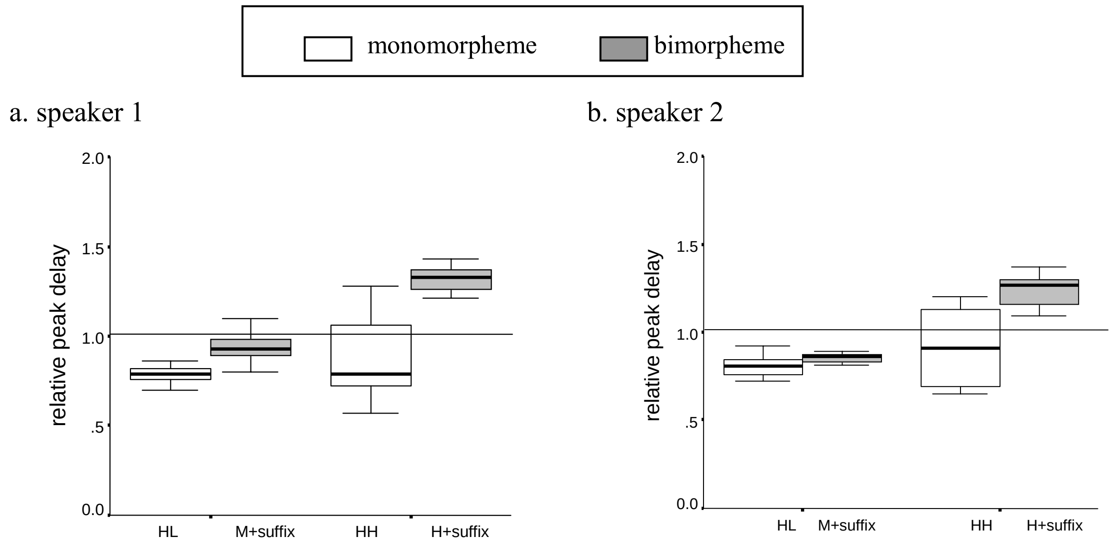


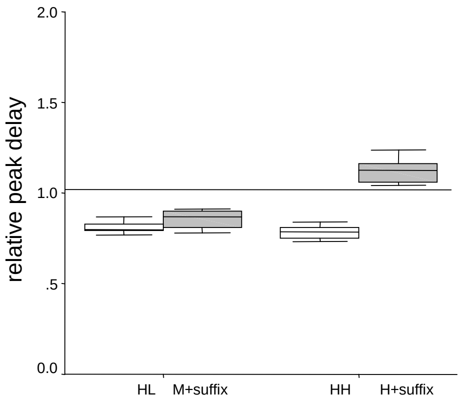

Figure 3.11: Typical effects of monomorphemic and bimorphemic words on the timing of F0

The peak is in the first syllable for (a) HL and (b) M+suffix. In contrast, the peak begins in the first syllable, but is extended onto the second syllable, showing the long peak plateau for (c) HH. However, the peak is in the second syllable for (d) H+ suffix. The unexpected pattern is that the peak occurs earlier for (a) HL than for (b) H+ suffix.

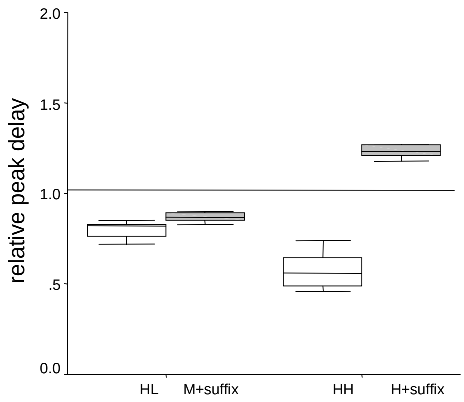
A quantitative measure of the difference in the timing of F0 peak is the *relative peak delay*. Fig. 3.12 shows the variation in *relative peak delay* in a box plot graph with tone type and morpheme type compared for each speaker. It is seen that the medians in HH are higher than in “H+suffix”, indicating that the peak came later in monomorphemic HH than in bimorphemic “H+ suffix”. Also the medians are “less than 1” for HH, but “greater than 1” for “H+suffix”. This indicates that the peak was *in* the first syllable for monomorphemic HH, but it was *after* the first syllable for bimorphemic “H+ suffix”. It definitely shows that the claim that the bimorphemic “H+ suffix” has the same realization as HH is not plausible.



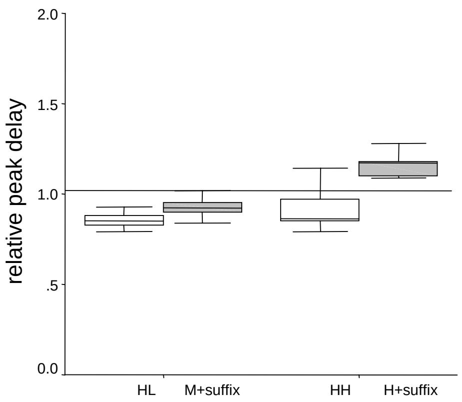
c. speaker 3




d. speaker 4




e. speaker 5




f. speaker 6

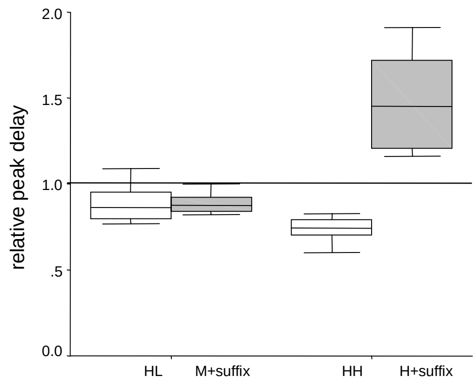



Figure 3.12: *Relative peak delay* for HL, M+suffix, HH, and H+suffix

The data were analyzed statistically by means of repeated measures ANOVA in which the dependent variable was the relative peak delay. Subject was the between-subjects factor, and tone type (tone type 1 and tone type 2) and morpheme type (monomorpheme, bimorpheme) were the within-subjects factors.

Although there were significant main effects of tone type ($F(1,54)=187$, $p <.001$) and morpheme type ($F(1,54)=158$, $p <.001$), these main effects are probably significant only due to the interaction of tone type and morpheme type, because the means are greater for “tone type 2” than “tone type 1”, not across the board, but only in bimorphemic words, i.e., “H+suffix”. The interaction of the tone type and morpheme type was significant ($F(1,54)=104$, $p <.001$). This might be because the peak usually occurs in the second syllable in “H+suffix” while the peak regularly occurs in the first syllable in other tone classes. A pair-wise comparison by a post hoc test showed that the mean relative peak delays were significantly greater for bimorphemic “H+suffix”. One unexpected result is that the relative peak delay was significantly greater for “M+suffix” than for HL, suggesting that the peak might be earlier in monomorphemic HL than in bimorphemic H+L, because the tone type “M+suffix” was suggested as a H+L in Experiment 3.1.

*3.2.2.3 The timing of fall*

We investigate the difference in the fall delay between bimorphemic “H+ suffix” and “M+suffix”, and monomorphemic HH and HL in this section. The hypothesis was that the fall is later in HH than in HL, because the peak occurs in the initial syllable for HL but both in the initial and second syllable for HH. A quantitative measure of the difference in the timing of F0 fall, *relative fall delay*, was measured for all four types. Fig. 3.13 shows the variation in relative fall delay in a box plot graph with tone type compared for each speaker.

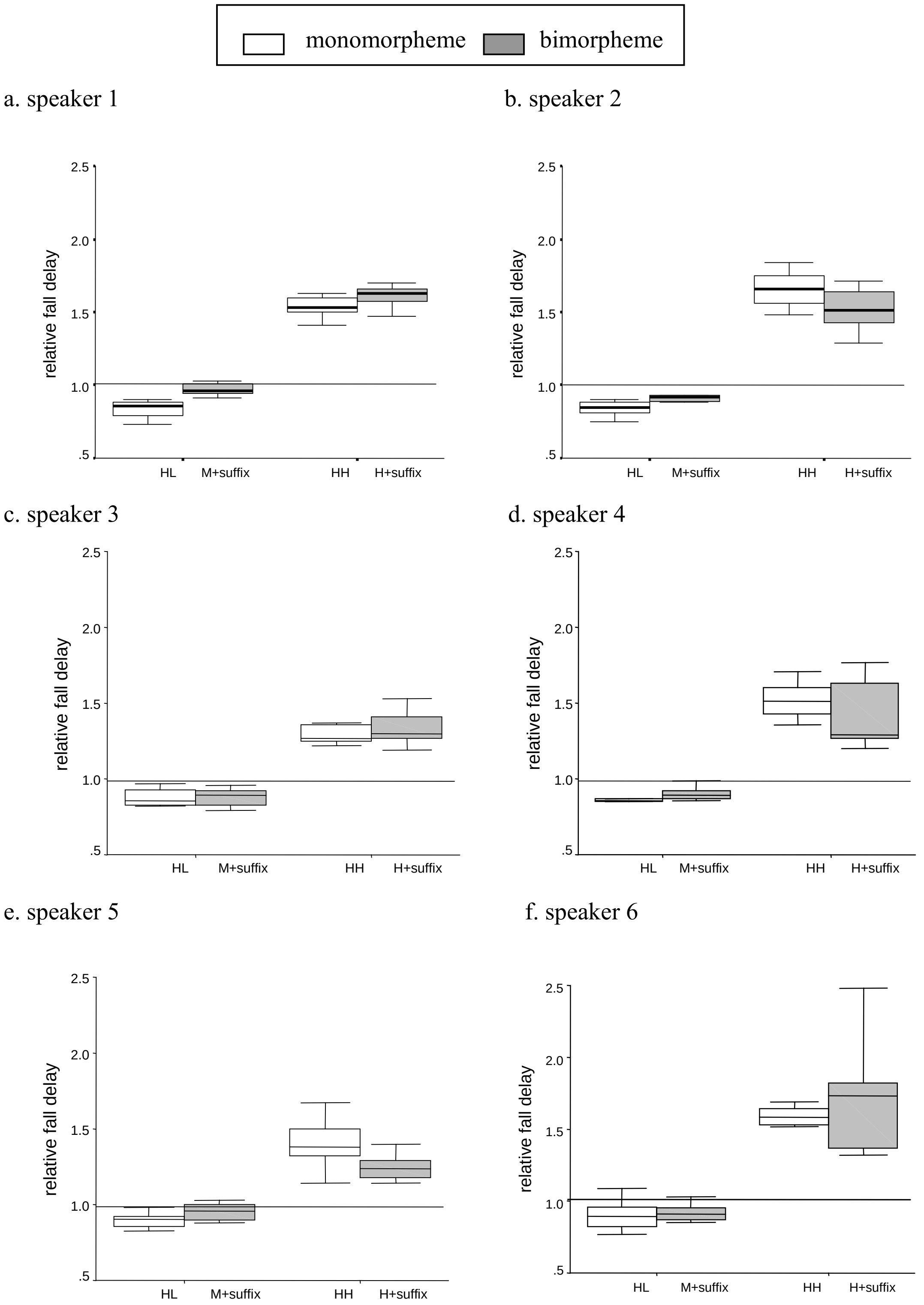



Figure 3.13: *Relative fall delay* for HL, M+suffix, HH, and H+suffix

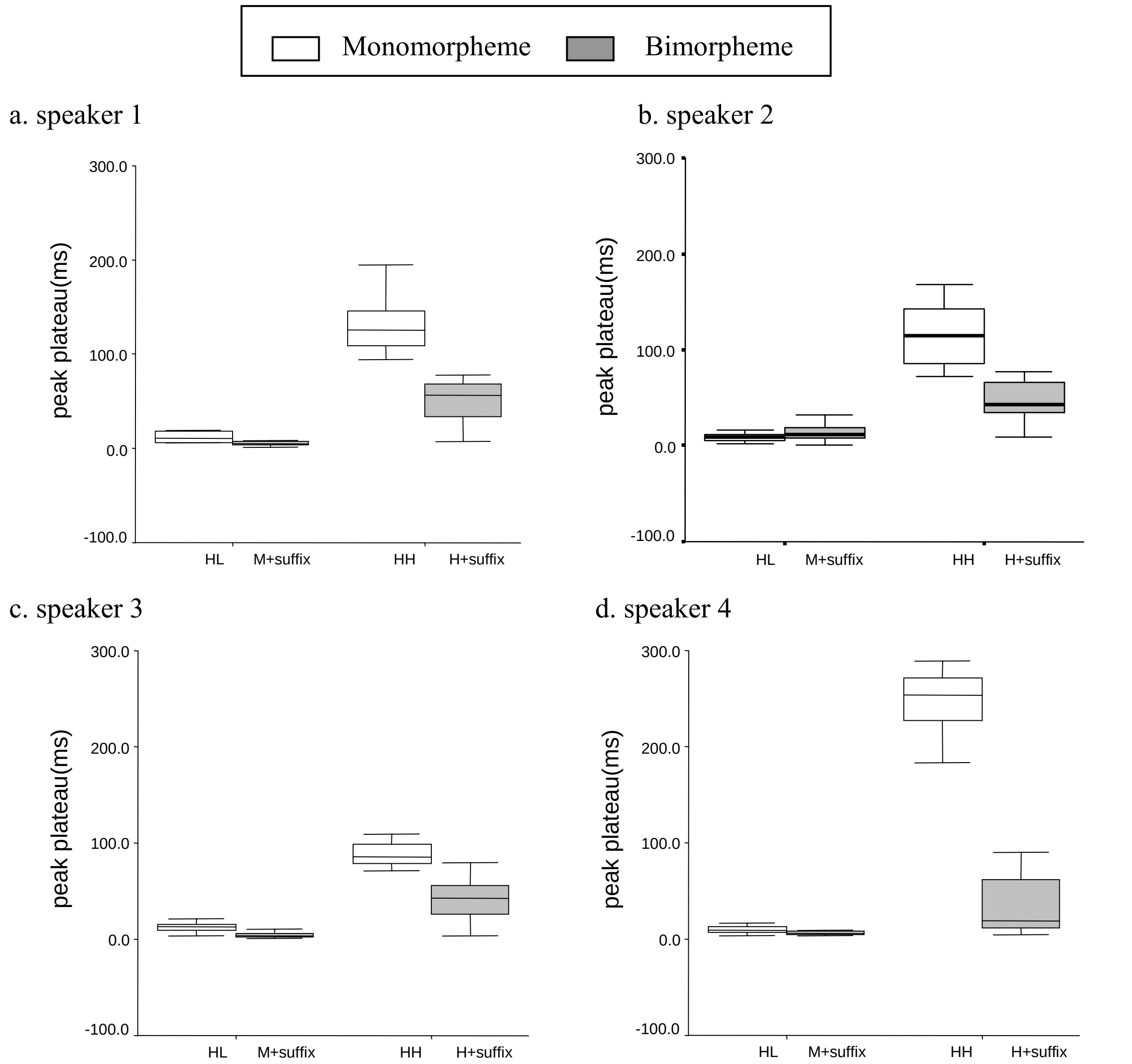
The medians are "less than 1" for first two boxes but "greater than 1" for last two boxes, indicating that the F0 fall was in the first syllable for HL and "M+suffix" but was after the first syllable for HH and "H+suffix".

These data were analyzed statistically by means of repeated measures ANOVA in which the dependent variable was the relative fall delay. Subject was the between-subjects factor, and tone type (tone type 1: HL, "M+suffix", tone type 2: HH, "H+suffix") and morpheme type (monomorpheme, bimorpheme) were the within-subjects factors.

There was a significant main effect of tone type ($F(1,54)=1452$, $p <.001$), indicating that the mean F0 fall delays are significantly greater for tone type 2 (HH, H+suffix). This supports the hypothesis that the relative fall delay is greater for HH than for HL. However, the significant main effect of morpheme type, and the interaction of tone and morpheme type were not found. A pair-wise comparison by a post-hoc test showed that the mean relative fall delays were significantly greater for tone type 2 (HH, H+suffix) than for tone type 1 (HL, M+suffix), confirming that the F0 fall comes in the first syllable in HL and "M+suffix", but comes in the second syllable in HH and "H+suffix".

#### *3.2.2.4 Peak plateau*

The peak plateau difference between monomorphemic HH and bimorphemic "H+ suffix", and between monomorphemic HL and HH is tested. The hypothesis was that the peak plateau is greater for HH than for "H+suffix" and HL. Fig. 3.14 shows the variation in *peak plateau* in a box plot graph with tone type and morpheme type compared for each speaker. The median is higher for HH than for "H+suffix", indicating that the peak plateau is greater for monomorphemic HH than for bimorphemic "H+ suffix". And the median is higher for HH than for HL, indicating that the peak plateau is greater for monomorphemic HH than for monomorphemic HL.



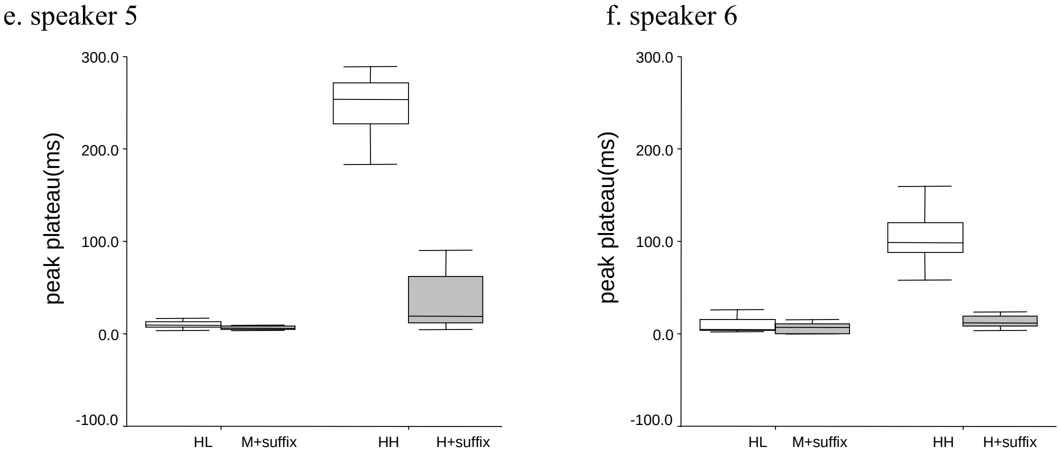



Figure 3.14: *Peak plateau* (ms) for HL, M+suffix, HH, and H+suffix

The data were analyzed statistically by means of repeated measures ANOVA in which the dependent variable was the peak plateau. Subject was the between-subjects factor, and tone type (tone type 1 and tone type2) and morpheme type (monomorpheme, bimorpheme) were the within-subjects factors.

There was a significant main effect of *tone* type ($F(1,54)=844$, $p <.001$), indicating that the mean relative fall delays are significantly greater for tone type2 (HH, H+suffix) than for tone type 1 (HL, M+suffix). The main effect of *morpheme* type was also significant ($F(1,54)=180$, $p <.001$), indicating that the mean relative fall delays are significantly greater for bimorphemic words than for monomorphemic words. The interaction of tone type and morpheme type was significant (($F(1,54)=182$, $p <.001$), with great monomorpheme effect for tone type 2 (HH). A pair-wise comparison by a post-hoc test showed that the mean peak plateau was significantly greater for HH than for "H+suffix", and greater for "H+suffix" than for the other classes, supporting the hypotheses.

### 3.2.3 Discussion

The results provide physical evidence for the differences among four tone classes, and the detailed differences of four tone classes can be expressed with the representation in Fig. 3.13. The suffix (*i* 'nominative', *imən* 'if') is underlined and italicized.

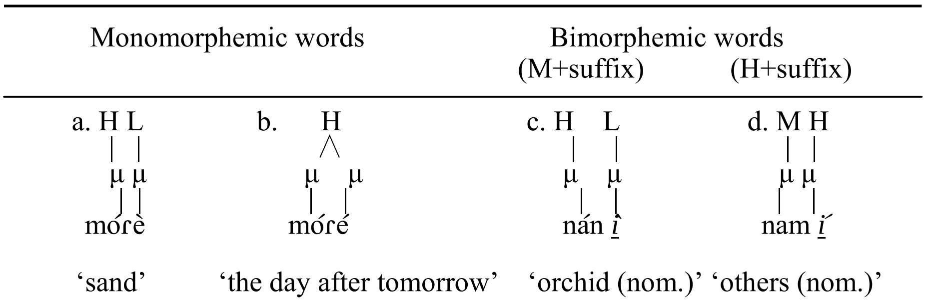


Figure 3.15: Representations of tone types for two morpheme types

The empirical support for the difference of two monomorphemic tone classes we call HL and HH can be summarized as follows. They are distinct in the timing of peak and fall. The F0 peak and fall was in the initial syllable for HL but the peak was stretched over two syllables for HH. These differences are represented in Fig. 3.13 that the high tone is associated with the first syllable for HL and with both syllables for HH, as seen in (a) and (b).

In addition, HL and HH are different in the peak F0 value for some speakers, that is, the peak was higher for HL than for HH. This lower peak pitch for HH was already noted in earlier study (Chung 1980). Due to this lower F0 peak of the tone class we call HH, this tone class might be described as MM in some of the impressionistic transcriptions (Kim 1980, Huh 1985). Similarly, there's some difference in peak F0 values among four tone classes, but consistent trend was not found. Therefore, it is suggested in Fig. 3.13 that the peaks in four tone classes are all represented to be associated with a high tone. The general pattern of the lower F0 in HH, or rather higher F0 in HL, can be understood to be a phenomenon which I would call anticipatory raising *or* H-raising. It has been found in quite a few languages. The higher F0 is raised by any following tone that has a low pitch value. This phenomenon has been reported for Thai by Gandour et al. (1994), for Mandarin by Xu (1997), Yoruba by Laniran & Clements (2003), and most recently, for Cantonese by Wong (2006).

The acoustic evidence of the representation H+L and M+H for the tone classes we call "M+suffix" and "H+suffix" was provided in Experiment 3.1. One of the main goals in this experiment was to compare the phonetic realization for the monomorphemic words and bimorphemic words that are described in a same way in earlier studies (Ramsey 1975, Kim 1996). Regarding this issue, two main differences were found relative to these previous studies.

The first is the difference between monomorphemic HH and bimorphemic H+suffix (M+H). While earlier studies have proposed the same tonal representation for these two tone types (Chung 1980, Kim 1996), the acoustic data provides the evidence against them. The peak plateau was significantly longer in HH than in H+ suffix (M+H). One possibility is that since two monomorphemic words used in this experiment, mə́ré (HH) and mə́re (HL), are tone minimal pairs, speakers tend to make a greater effort in order to contrast these two lexical items than they do with bimorphemic words. However, this cannot explain why the peak was generally realized in the suffix (the second syllable) and not in the root (the first syllable) in bimorphemic H+ suffix (M+H).

One more probable phonetic explanation might be the effect of intrinsic vowel height on F0: high vowels have a higher $F_0$ than do low vowels (Peters & Barney 1952, Lehiste & Peterson 1961, Mohr 1971, Hombert, Ohala & Ewan 1979). The vowel /i/ in the second syllable has higher intrinsic F0 than does /a/ in the first syllable of the test syllable, na + mi '*other people, nom'*. This could make the second H slightly higher than the first H, and thus the peak is realized only in the second syllable. However, it was seen in my informal work that the sample pitch track for bimorphemic H+ suffix, i.e., nu.ni '*eye (nom.)*', also has the peak in the second syllable, although the vowels in the two syllables are same high vowels.

Therefore, it is argued that the data in this experiment do not verify the claim in literature that the bimorphemic H+suffix and monomorphemic HH have same tonal representation (Ramsey 1975, Kim 1996). The data suggested the morphological structure may be encoded in the phonetic realization of F0.

Moreover, the difference in F0 realization was found for two tone classes we call HL and M+suffix (H+L), which are described in a same representation (Ramsey 1975, Kim 1996). They are different in the peak F0 for some speakers and in the timing of the peak. The peak came earlier for HL than for M+suffix (H+L), and the peak F0 values were higher for HL than for M+suffix (H+L).

### 3.3. EXPERIMENT 2C: SINGLE H AND DOUBLE H CONTRAST IN PERCEPTION

In the previous section, we saw that F0 fall delay and peak plateau were the main factors to distinguish the monomorphemic single H (HL) from the doubly-linked H (HH), and thus we can tentatively hypothesize that native speakers of South Kyungsang Korean make use of the those factors in perceiving the single H and double H in one-morpheme words, and there are definitely tonal contrasts of HL and HH. However, this issue has not been investigated to my knowledge. In this section, I examine whether the peak plateau indeed contributes to the perception of single H and double H contrast. Since varying peak plateau automatically also varies F0 fall delay, I consider only peak plateau as a factor.

The experiment tests the perception of isolated synthetic stimuli in which the peak plateau has been systematically manipulated relying on the production study. Since the peak plateau was significantly longer for HH than for HL in acoustic data, the prediction is that the stimuli would trigger HH tone responses when the peak plateau is longer, and the stimuli would trigger HL tone responses when the F0 peak plateau is shorter. The response time is also considered. The stimuli with extreme values of parameters would trigger the fast response time, while stimuli with between values would trigger the relatively slow response time.

#### 3.3.1 Methods

##### *3.3.1.1 Stimuli*

To test the hypothesis, two words that differ minimally in lexical tone, /moɾe/ (HL ‘sand’ – HH ‘the day after tomorrow’), were chosen. One female native speaker of South Kyungsang Korean produced the isolated HL /HH tone minimal pairs: /móɾe/ (HL) ‘sand’ - /móɾé/ (HH) ‘the day after tomorrow’, at the normal rate of speech. Recordings were made in a sound-treated booth in the Phonetics lab of the Linguistics department, University of Texas at Austin, using Praat.

With the original sound object, I created a / moɾe / continuum using Praat on a PC, by removing the pitch points except for three points, i.e., initial, peak, and final pitch point as in Fig.2.5. Since the F0 contours in manipulation window of Praat consists of a number of pitch points, the pitch contour is needed to be stylized only with the relevant pitch points. This procedure was done with the cutting function in Praat. Then, I manipulated each pitch point according to the proposed values, by dragging the pitch point, and generated 32 different continuums.

The F0 values and durations for four points in the original recording were 220Hz (0 ms), 275Hz (136 ms), 275Hz (156 ms), and 154Hz (530 ms), and the syllable duration was 530(ms). The peak plateau in original recording was 20ms, so I manipulated the peak plateau, by increasing the peak offset point in step of 40 (ms) from the original peak plateau duration (20ms) as in Fig. 3.13, generating 8 continuums.

Therefore, the /mớɾe/ - /mớɾé/ continuum consisted of eight /moɾe/ syllables, whose F0 values at the onset, peak and offset points were 220Hz, 275Hz, and 154Hz, as in the original values, and the duration of the syllables was also fixed at 530ms. The only varying parameter along the continuum was the peak plateau duration (20ms, 60ms, 100ms . . . 300 ms). Although only peak plateau was intended to manipulate in the experiment, the slope was also varied, because the syllable duration was fixed at constant value in this experiment and thus the slope was getting steeper as the peak plateau was increased. Fig. 3.16 shows the F0 contours of the eight syllables, from the syllable onset to the syllable offset.

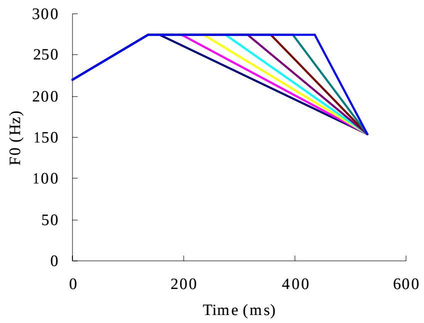


Figure 3.16: F0 contours of the /moɾe/ continuum

#### *3.3.1.2 Subjects*

10 native speakers of South Kyungsang Korean participated in the experiment, all of whom had also participated in the preceding Experiment 1b.

#### *3.3.1.3 Procedures*

The stimuli were played out 10 times in a random order on an audio system using an ALVIN software program on a PC. So each subject heard 80 stimuli (8 peak plateau * 10 repetitions). There was a two-second interval between stimuli. The subjects were instructed to respond to each item as quickly as possible by pressing the button corresponding to English words and pictures for “sand” (HL) or “the day after tomorrow” (HH) because Korean orthography is almost same for two /moɾe/. The used pictures for each response button were given in Fig. 3.17.

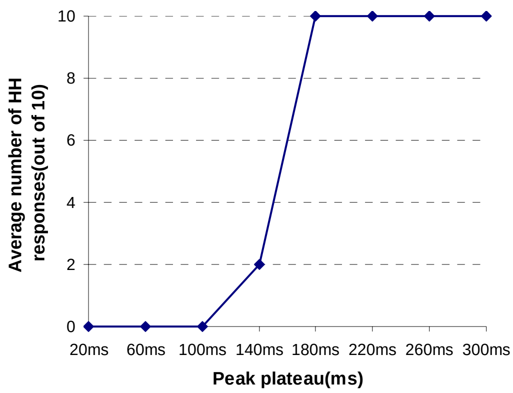
Figure 3.17: Pictures for response button used in perception experiment of HL and HH

To avoid misidentification, subjects were asked to speak two minimal pairs, by pronouncing the corresponding word when I pointed the button label, prior to the experiment. A practice session consisting of 10 test items in each block preceded the test as well. These practice items provided listeners with end points for each parameter manipulated, as well as stimuli between. Subject responses were collected by computer using an ALVIN program, and responses for each stimulus were added across speakers.

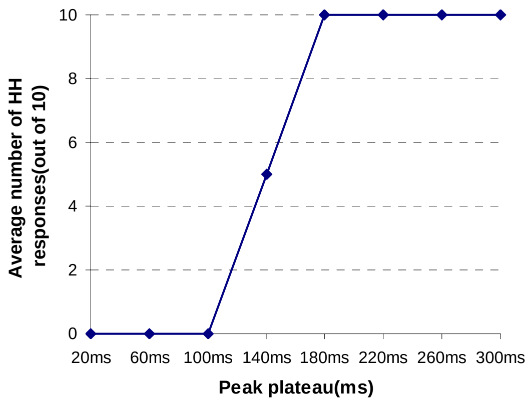
### 3.3.2 Results

The identification results from the ten subjects are consistent with the hypothesis. The stimuli which have a longer peak plateau are more likely to be heard as HH; the stimuli which have a shorter peak plateau are less likely to be heard as HH. The Fig. 3.18 gives the average number of HH -/mə́ɾé/- responses at each stimulus for each subject. In the Fig. 3.18, the y-axis indicates the "*total number of HH response*" out of 10, and x-axis indicates the 8 variants of peak plateau (20ms, 60ms, 100ms, 140ms, 180ms, 220ms, 260ms, and 300 ms). Although there were small differences in the boundary of sharp curve for each subject, the figure generally shows that the rise is parallel, beginning at 4$^{th}$ point, i.e.140 (ms) and reaching peak slope at 5$^{th}$ point, i.e., 180(ms), showing the sharp rising curve.

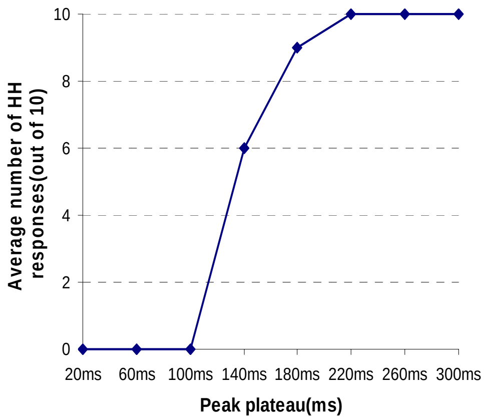
a. subject 1



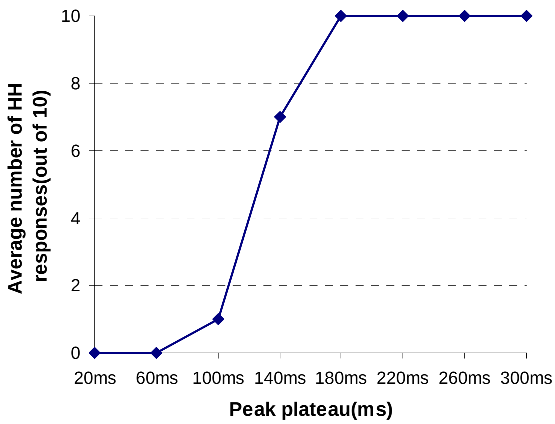

b. subject 2



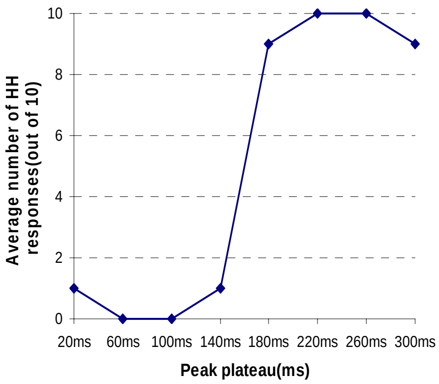

c. subject 3



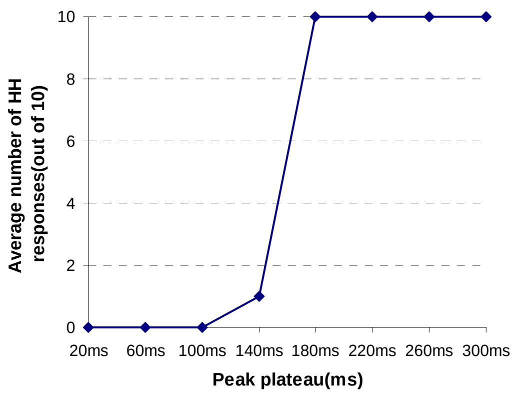

d. subject 4



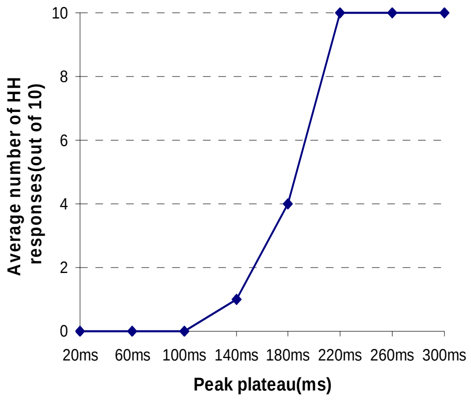

e. subject 5



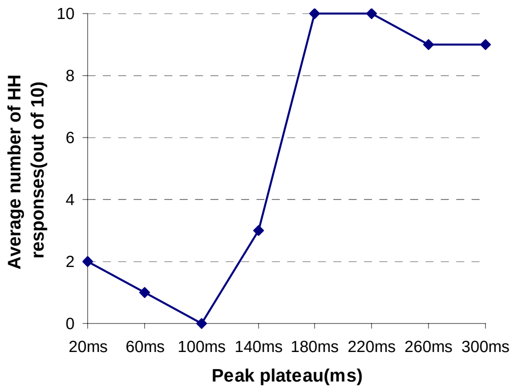

f. subject 6



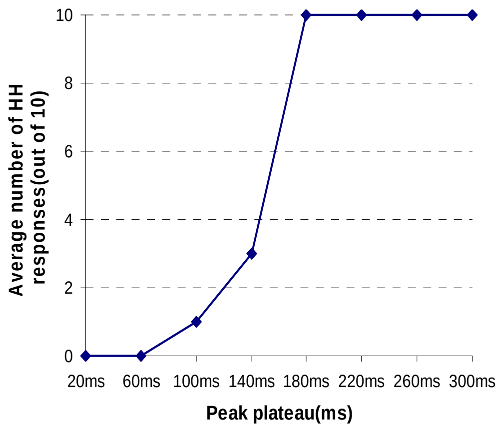
g. subject 7




h. subject 8




i. subject 9




j. subject 10

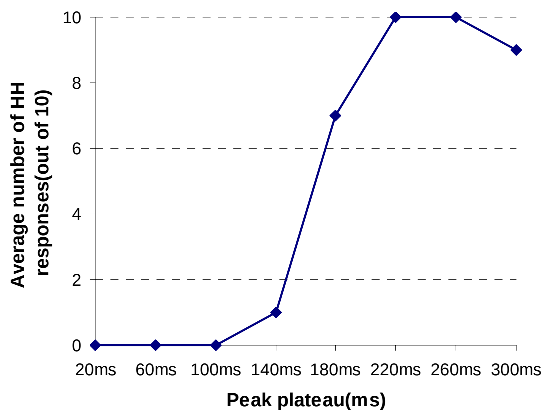



Figure 3.18: The total number of HH responses for each subject

To determine whether the factor -peak plateau- contributes to distinguish the HH from HL, a logistic regression analysis of the whole dataset pooled across subjects was used. The independent variable was the peak plateau (in milliseconds), and the dependent variable was the dichotic answer choice (HL or HH). HL was coded as 0 and HH as 1.

The equation of such model is presented in Table 3.3, together with the coefficients (and odds ratios) and p-value.

| a (peak plateau) | b constant | p |
|---|---|---|
| .049 (1.050) | -7.385 (.00) | p<.001 |

*Equation*: Probability of HH response = (a × peak plateau) + b

Table 3.3: Logistic regression analysis

The statistical results support the hypothesis. The result shows that the peak plateau factor was significant (p<.001) in the logistic regression analyses. The positive coefficient indicates a direct relationship to the dependent variable. That is, as the peak plateau is longer, there is a higher likelihood for the HH response. The coefficient, for example, can be interpreted that 1 (ms) longer peak plateau is 1.05 times more likely to be judged as an "HH".

### 3.3.3 Discussion

The goal of experiment 2c was to determine whether the acoustic dimension of peak plateau contributes to perception of monomorphemic HL and HH in South Kyungsang Korean. The results of the logistic regression analysis seen in Table 3.21 can be interpreted as perceptual models of HH. It was seen that, for HH to be identified, the perceived target for the peak plateau must be long. This result confirmed the previous suggestion that there is a monomorphemic HL and HH contrast (Huh 1955, Ramsey 1975, Kim 1996), by showing that the listeners of South Kyungsang Korean make use of the peak plateau factor in perceiving the single H and doubly-linked H in one-morpheme words.

Furthermore, it was shown that the boundary between the two categories in the perception data was consistent with the acoustic data, i.e., the mean peak plateau of HH for 6 speakers in the acoustic data was 146 (ms) (*see* Table 3.11, p.116), and the sharp binary distinction in perception data was between 140(ms) -180(ms), corresponding the 4th and 5th stimulus in general. A peak plateau longer than 140(ms) tends to be judged as an "HH" tone, while a peak plateau less than 140(ms) is judged as an "HL" tone. The boundary for the response time was also consistent with the data of response time, i.e., the response time had a peak at 140(ms), suggesting that the stimuli were heard as most ambiguous during this peak plateau duration.

The conclusions drawn from these data are that, in South Kyungsang Korean, F0 peak plateau was a perceptual cue to trigger categorical identification functions between a monomorphemic single H and a doubly-linked H, and there is an HL and HH contrast in this language. However, the stimuli used in this test differ not only in the duration of the peak plateau, but also in the slope of the F0 fall and relative fall delay (*see* Fig. 3.13, p.124), the schematization of the /moɾe/ continuum. This study cannot determine whether these slope and relative fall delay differences also play a role in triggering the listener's responses. Further studies are needed to explore this possibility.

# Chapter 4: Conclusion

This research increases our understanding of South Kyungsang Korean tonology, suggesting explanations for why tonal descriptions have been inconsistent among authors and for how tone alternation patterns in suffixed words have a significant function in distinguishing the tone classes in this language.

The data suggest that the South Kyungsang Korean has three tone targets: high (H), mid (M), and low (L), and in each word, there is one and only one sequence of F0 peak, consisting, at most, of two syllables. The surface representations of tone classes were proposed and are repeated in Fig. 4.1. The suffix (*i* 'nominative', *imən* 'if') is underlined and italicized.

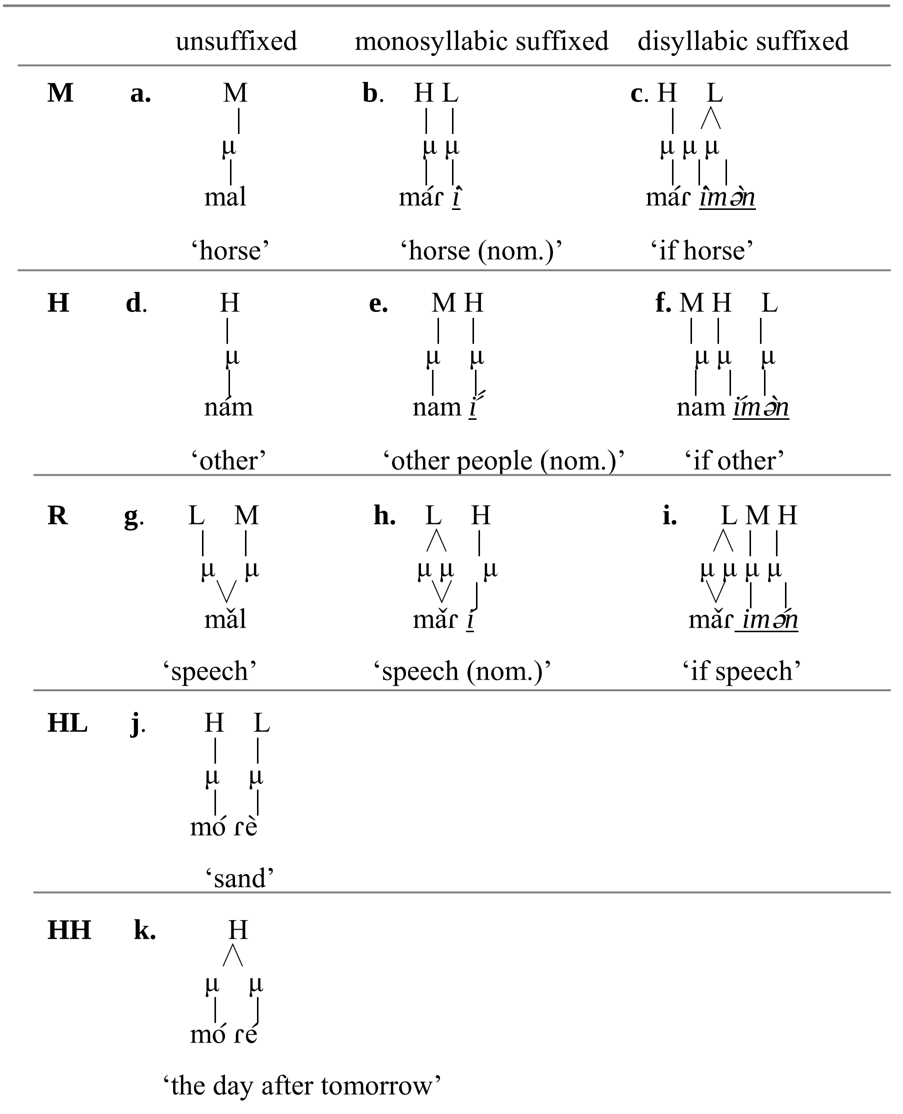



Figure 4.1: Representations of tone types in South Kyungsang Korean

The empirical support for these tone classes was provided through a set of production and perception experiments, yielding four major findings. First, in unsuffixed words, the tone class we call M has a peak pitch that is slightly lower than the peak in tone class we call H, but there’s no consistent difference in peak F0 values between M and R. Instead, the initial F0 values were significantly lower for R than for M. The syllable duration of M is slightly longer than H but is noticeably shorter than R. In addition, the peak occurs later for R than for M, and there's no substantial difference in peak delay between M and H. These differences in unsuffixed words are represented in Fig. 4.1 that the mora is associated with a mid tone for M-class and associated with a high tone for H-class. On the other hand, two moras are employed to represent the significantly longer syllable duration and later peak for R-class than for others. The initial mora is associated with a low tone and the second mora is associated with a mid tone for R-class.

Second, when followed by a monosyllabic suffix or vowel-initial polysyllabic suffix, the peak regularly occurs in the root for M-class, and the peak F0 values were higher in unsuffixed M-class. This is represented that the first mora is associated with a high tone for M-class in suffixed words.

Third, when followed by a monosyllabic suffix or vowel-initial polysyllabic suffix, the peak occurs in the early portion of the initial syllable of the suffix for H-class. Since the peak is realized with F0 rising through the initial syllable up to the onset of the second syllable, both the initial and the second syllables have high F0 values. Therefore, the peak is, at times, stretched over two syllables, but usually occurs only in the second syllable. This pattern is distinct from the one observed in monomorphemic HH in which the peak is regularly extended over two syllables. This is illustrated in the Fig. 4.1 that the first mora is associated with a mid tone and the second mora is associated with a high tone.

Fourth, when followed by a monosyllabic suffix or vowel-initial polysyllabic suffix, the root has a low pitch, and the peak occurs in the middle portion of the suffix in monosyllabic suffixed words. In disyllabic suffixed words, the peak occurs in the early portion of the final syllable of the suffix. The first syllable durations in suffixed words of R-class are also longer than that of M-class. This are illustrated that the first syllable is associated with two moras for R-class, and the initial syllable is associated with a low tone and the second syllable is associated with a high tone in monosyllabic suffixed words. Similarly, the third syllable is associated with a high tone in disyllabic suffixed words, and the F0 rising contours of intermediate moras between low toned and high toned ones are assumed to be a mid tone although their F0 values were not measured in this work.

In addition, acoustic evidence for two tone classes in monomorphemic disyllabic words (HL and HH) was provided, compared with the bimorphemic disyllabic words. The tone classes we call monomorphemic HL and bimorphemic H+L (M+suffix) have a peak in the first syllable and thus same surface representation was suggested for them, as in Fig. 4.1 (b and j), but the fine-grained phonetic data showed that the peak comes earlier in HL than in H+L (M+suffix). The tone classes that we call monomorphemic HH and bimorphemic M+H (H+suffix) also different each other, that is, HH has a long peak plateau stretched over two syllables, but this was not regularly observed in bimorphemic M+H (H+suffix).

Consequently, the findings of this study suggest that the basic facts of South Kyungsang tonology need to be revised. Although the results are consistent with the claim that there are three distinct tone classes (Choi 1929, Huh 1955, Ramsey 1975, Kim 1996), they do not support either claim that the tone class we call M is a high-tone and the class we call H is a mid-tone (Huh 1955, Kim 1974) *or* the claim that M and H are the same high tone (Cho 1996, Kim 1996). Rather, the peak F0 values were higher in H-class than in M-class and the difference was statistically significant for some speakers.

The data also do not verify the claim that the tone class we call R is a low tone (Huh 1955, Kim 1974) *or* that R is a high tone (Kim 1996). Although the initial F0 values of R were low, it has a later peak that is high as in M-class.

Further, the findings do not prove the argument that a high tone is associated with two syllables when H-class and R-class are followed by a suffix (Kim 1996). The peak is realized with gradual F0 rising through the two syllables, and thus the peak is, at times, stretched over two syllables for some speakers, but it was not regular, even within one speaker. The peak was usually came after the syllable bearing H-class, *or* came in the second syllable, following from the syllable bearing R-class.

Moreover, they provide no support for the claim that M and H with a monosyllabic suffix (M+suffix, H+suffix) have the same tones with the monomorphemic HL and HH (Kim 1996), as noted above.

These results serve as another way of understanding tone targets and tone alternation patterns in this language. I suggest that the different tone alternation patterns in three tone classes are the reflection of the phonetic implementation of three tone targets. For example, the peak is in the final portion of a syllable in R-class, and it may take more time for the peak to be fully realized, compared to the simple tone target. This leads to long peak delays for R-class and thus the F0 peak is realized in the second syllable, following from the item in R-toned syllable in disyllabic suffixed words (Chen & Xu 2006). Likewise, it would take longer time to reach a high peak (H-class) than a lower peak (M-class), and this might bring about longer peak delays for H-class than for M-class. It was seen that the peak occurs after the syllable that carries the item in H-class in suffixed words (Silverman & Pierrehumbert 1990, Prieto et al., 1995, Myers 1999, Xu 2001, 2006). Given that the difference in peak F0 values between H-class and M-class were not great in unsuffixed words, the peak delay effect might not be stimulated in unsuffixed words. However, this effect can be exploited in suffixed words in which the peak can be fully implemented.

The lack of long peak delays in unsuffixed words (word-final syllable) is also well reported (Silverman & Pierrehumbert 1990, Prieto et al., 1995). When a syllable is lengthened, due to prosodic factors, such as an upcoming stress, pitch accent or phrase boundary, then, the F0 peak moves closer to the syllable onset.

Regarding the lack of the peak delay in monomorphemic words, it is suggested that the peak might be pushed to the onset of the syllable in monomorphemic words because the following lexical tone should be implemented immediately, while there would be no such hurry when the following tone is weak, as in suffixed words. Given that this language has a phonemic contrast between early and late peaks (HL vs LH), if the peak in HL is delayed to the following syllable, the lexical contrast would be ambiguous. This lack of long peak delay was also reported in other languages that has a lexical contrast between early and late peaks, such as Kinyarwanda (Myers 2003) and Swedish (Bruce 1977).

This work suggests a number of possibilities for future research. The production experiments using F0 timing measurements and perception experiments using an identification task can shed light on the research that investigates the phonological categories of tones. Specifically, the investigation of F0 timing on tone alternation pattern can be useful in other languages for determining tonal targets. As a parallel, we can find a significant clue for the question regarding the nature of tone in long vowels of North Kyungsang Korean, which was the research question in the pilot studies of the author (Chang 2003, 2005). When a long vowel word is followed by a disyllabic suffix (e.g., pə́da 'than'), the high tone occurs in the final syllable of the suffix for South Kyungsang Korean. By contrast, the peak occurs in the initial syllable of the suffix for North Kyungsang Korean, as in (22). The example for North Kyungsang Korean is from Kim 1997.

(22) a. <u>South Kyungsang Korean</u>

H: mul-*póda* 'than water'

R: mà:l-*podá* 'than speech'

b. <u>North Kyungsang Korean</u>

H: múl-*póda* 'than water'

H: <u>ma:l</u>-*póda* 'than speech'

Although the tonal descriptions of the underlined long vowel of North Kyungsang Korean vary among authors (Chung 1991, Kim 1997), the authors agree that the first syllable of the suffix has a high tone. None of them argue that the final syllable of the suffix has a high tone. This alternation pattern weighs for the claim that the long vowel in question is a same tone with the short vowel. If the long vowel in North Kyungsang Korean is a different tone, the peak might be more delayed compared to a short vowel, as seen in South Kyungsang Korean, *or* at least the tone alternation pattern for a long vowel should not be the same with the pattern for a short vowel. However, the tone alternation pattern is identical for a short vowel and a long vowel for North Kyungsang Korean, suggesting that two length variations might be the same tone.

Moreover, this work opens the possibility that morpheme structure can be encoded in the phonetic realization of F0 contours. First, it was shown that the long peak delay was not found in monomorphemic HL and HH. Instead, the F0 peak for HL occurred quite earlier than for a bimorphemic H+L (M+suffix). Second, the peak was higher for HL than for H+L (M+suffix) for some speakers. While there has been research on morpheme effects on phonetic realization of segments (Cho 2001), few studies have addressed the effects on suprasegmentals such as tones. This approach can inspire future research that explores further the F0 realizational difference correlating with the morphemic structure of the words. One possible follow-up study is to examine the phonetic realization of F0 contours for the monomorphemic LH (e.g., àngé 'fog') and bimorphemic L+H (e.g., mà+rí ‘speech (nom.)’) in this language, which was not investigated in the current work. This comparison will provide additional information about the F0 realizational difference correlating with the morphemic structure of the phonological words.

The perception experiments in this study were pioneering work for this particular dialect of Korean. One pattern worth further study is the great between-subject variations in perception task. Subjects were all able to distinguish the tone categories, but used different weightings of the cues to identify the word. It was proposed that if the speaker and listener are the same gender, the listeners may be more familiar with the talker’s voice, and thus perceive the stimuli more precisely than listeners from the other gender which would result in less *within*-subject variation. Further, a correlation of the variations in the perception task with variations in the production task was suggested. For example, the speaker with a markedly lower initial F0 for R in the production test showed a stronger cue for the low initial F0 factor in the perception test. It was unfortunate that the subjects participating in the production test were not completely matched with the subjects participating in the perception test, and thus a thorough comparison cannot be made. As there has been little discussion of variability among speakers in perception, attention to more details with more data about this issue will deepen our understanding for the relationship of production and perception research.

# References

Abramson, A.S. 1962. The vowels and Tones of Standard Thai: Acoustical Measurements and Experiments. *Publication 20, Indiana University Research Center in Anthropology, folklore, and Linguistics*, Bloomington.

1975. The tones of Central Thai: some perceptual experiments. In J.G. Harris and J. Chamberlain (eds.), *Studies in Tai Linguistics*. Bangkok: Central Institute of English Language, p.1-16.

1979. Lexical tone and sentence prosody in Thai. In E. Fischer-Jorgensen, J. Rischel, and N.Thorsen, eds, *Proceedings of the Ninth International Congress of Phonetic Sciences* 2:380-387. Copenhagen:University of Copenhagen.

Archangeli, D. 1984. Underspecification in Yawelmani phonology and morphology. PhD dissertation, MIT.

Arvaniti, A., D. R. Ladd, and I. Mennen. 1998. Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics* 26:3-25.

Beckman, M.E., and J. Pierrehumbert 1986. Intonational Structure in Japanese and English. *Phonology Yearbook* 3:255-309

Browman, C and L. Goldstein. 1986. Toward an articulatory phonology. Phonology Yearbook 3: 219-252

Brown, G., K.L. Currie, and J. Kenworthy. 1980. Questions of intonation. London: Croom Helm.

Bruce, G 1977. *Swedish word accents in sentence perspective*. Lund:Gleerup.

Butcher, A. and J. Harrington. 2003. *An instrumental analysis of focus and juncture in Warlpiri*. 15th ICPhS Barcelona.

Caspers, J and V.J. van Heuven. 1993. Effects of time pressure on the phonetic realization of the Dutch accent –leading pitch rise and fall. *Phonetica* 50:161-171.

Chang, S-E. 2003. Tones in North Kyungsang Korean. Unpublished.

2005. F0 timing in North Kyungsang Korean. *2005 Annual Meeting of Linguistic Society of America* (LSA), San Francisco, January, 6-9, 2005.

Chao, Y.R. 1968. *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.

Chen, Y. and Xu, Y. 2006. Production of weak elements in speech - Evidence from f0 patterns of neutral tone in standard Chinese. *Phonetica* **63**: 47-75.

Cho, T-H. 2001. Effects of morpheme boundaries on intergestural timing: evidence from Korean. *Phonetica* 58:129-162.

Cho, Y-H. 1996. Quantitative and prosodic representation of tone and intonation in the Kyungnam dialect of Korean. PhD dissertation, University of Arizona.

Choi, H-B. 1929. *Wuli Malpon* (Our grammar). Seoul: Yonsei University Press.

Chung, K. 1980. Neutralization in Korean: A functional view. PhD dissertation, University of Texas at Austin.

Chung, Y-C. 1960. Sibo seyki kukeuy Tone ey tayhan Yenku (Study on Tones in 15C), *Kuke Yenku* 8.

1974. Kyengsangdo Pangenuy Sengco Yenku (Study on Tones in Kyungsang Tone). PhD dissertation. Seoul National University.

1974. *Studies on the Tone of Kyungsang Dialect*, Seoul: Top Press.

1986. Kyungsangdo Pangen Sungco Yengu (Studies on the Tone of Korean), *Kukehakhwe*.

Chung, Y-H. 1991. The Lexical Tone System of North Kyungsang Korean. PhD dissertation, Ohio State University.

2002. Contour tone in the North Kyungsang dialect: Evidence for its existence. *Studies in Phonetics, Phonology and Morphology* 8(1): 135-147.

Delattre, P. 1966. A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics* 4:183-198.

Dilley , L.C., D.R. Ladd, and A. Schepman. 2005. Alignment of L and H in bitonal pitch accents" testing two hypotheses. *Journal of phonetics* 33:115-119.

D'Imperio, M. 2001. Focus and tonal structure in Neapolitan Italian. *Speech Communication* 33: 339-356.

Gandour, J.S. 1978. The perception of tone, In V.A. Fromkin (ed). *Tone: A Linguistic Survey*, p.41-76. New York: Academic press.

Gandour, J., S. Potisuk, and S. Dechongkit. 1994. Tonal coarticulation in Thai. *Journal of Phonetics*, 22: 477-492.

Guessenhoven, C. 1984. On the grammar and semantics of sentence accents. Foris: Cinnaminson.

Goldsmith, J. 1976. Autosegmental phonology. PhD dissertation, MIT.

Han, S.M., and K.O. Kim. 1974. Phonetic variation of Vietnamese tones in disyllabic utterances. *Journal of Phonetics* 2:223-232.

Hombert, J. 1978. Consonant types, vowel quality, and tone. In V. Fromkin ed., *Tone: a linguistic survey*. New York: Academic Press.

Hombert, J.-M., J. J. Ohala and W. G. Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55, 37-58.

Hong, Y. 1988. A Sociolinguistic Study of Seoul Korean. PhD dissertation. University of Pennsylvania.

Huh, W. 1955. Kyengsang do Pangen uy Sengco (Tone in Kyungsang dialect). *Festschrif for Prof. H.B. Choi,* pp. 447-519.

1972. Cungsey Kuke Yenku (A study of Middle Korean). Seoul: Cengumsa.

1985. Kyengsang do Pangen uy Sengco (Tone in Kyungsang Dialect), *Korean Journal* 25: 19-32.

Hyman, L. 1993. Register Tones and Tonal Geometry. In H.van der Hulst and K.Snider, eds. *The Phonology of Tone: The Representation of Tonal Register*. Mouton de Gruyter, Berlin, 75-108.

Hyman, L., and R. Schuh. 1974. Universals of tone rules. *Linguistic Inquiry* 5: 81-115.

Jun, S-A. 1993. The Phonetics and Phonology of Korean Prosody. PhD dissertation, Ohio State University.

Jun, J., J. Kim, H. Lee, and S-A Jun. 2004. The prosodic structure of Northern Kyungsang Korean. The proceedings of Speech Prosody, pp. 337-340.

Kang, K. 1986. Kyungnam Pangeun Sungco Yengu (The study of Tone in the Kyungnam Dialect), *Kyungsang Sungcoron*, 89-99, Hongmungak: Seoul

Kanerva, Jonni. 1989. *Focus and Phrasing in Chichewa Phonology*. New York: Garland.

Kenstowicz, M., and H.-S. Sohn. 1998. Accentual Adaptation in North Kyungsang Korean. *ROA*. 345-0999.

Kenstowicz & Park 2006. Laryngeal features and tone in Kyungsang Korean: A phonetic study. International Conference on East Asian Linguistics

Kim, C-K. 1973. Kuke Sengcolon kwa Sebu Kyengnam Pangenuy Sengco (Korean Tonology and the Tone in the South Kyungsang dialect). *Hankul* 152, 75-115.

1975. Kyengsangdo Pangenuy Sengcohyeng (The tonal type of the Kyungsang Dialect), *Ehak Yenku* 11.

1980. *Kyengsangdo Pangenuy SengcoChekye* (Tonal system in Kyungsang Dialect). Seoul: Kwahaksa

2003. *Yengnsam Pangen Sengco Pikyo* (Comparisons of tones in Kyungsang dialects). Seoul: Yuklak

Kim, G-R. 1988. The pitch accent system of the Taegu dialect of Korean with emphasis on Tone Sandhi at the Phrase Level. PhD dissertation, University of Hawaii.

Kim, M.-J. 1996. The tonal system of accentual language. PhD dissertation, University of Chicago.

Kim, N.-J. 1997. Tone, Segment, and their Interaction in North Kyungsang Korean: A correspondence theoretic account. PhD dissertation, Ohio State University.

Kim, W-C. 1974. A study on Middle Korean tones. PhD dissertation. Seoul National University.

Kim, Y-M. 1986. Kuke Puncelumsou Sacek Yenku (Historical Study of Segmental Phonemes in Korean), PhD dissertation. Korea University.

Kim, Y-M. 1986. Kyungnam Panenu Sungco Yenku (The study of Tone in the Kyungnam Dialect), *Enehak Yengu Nonmuncip, Kyungsangdo Pyun*, 3-3.

Kiparsky, P. 1982a. From cyclic phonology to lexical phonology. *The structure of phonological representations*, ed. By H. van der Hulst & N. Smith 1:131-175. Dordrecht:Foris.

1982b. Lexical morphology and phonology. *Linguistics in the morning calm*, ed. By I.S.Yang, 3-91. Seoul: Hanshin.

Klatt, D.1976. Linguistic uses of segment duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America,* 59, 1208-1221.

Kohler, K. 1987. Categorical pitch perception. In Arvo Eek (ed.), *Proceedings of the 11th International Congress of the Phonetic Sciences, Tallinn*, 5: 331-333.

Ladd, D. R., and K. E. A. Silverman. 1984. Vowel intrinsic pitch in connected speech. *Phonetica* 41: 31-40.

Ladd, D.R., and A. Schepman. 2003. "Sagging transitions" between high pitch accents in English experimental evidence. *Journal of phonetics* 31:81-112.

Laniran, Y. 1992. Intonation in tone languages: the phonetic implementation of tones in Yoruba. PhD dissertation. Cornell University.

Laniran, Y., and G.N. Clements. 2003. Downstep and high raising: Interacting factors in Yoruba tone production. *Journal of Phonetics* 31: 203-250.

Labov, W. 1994. *Principles of Linguistic Changes: Internal Factors*. Blackwell: Oxford.

Lee, H-B. 1989. *Korean grammar*. Oxford:Oxford University Press.

Lee, K-M. 1960. Sohakenhay ey tayhaye (On Sohak Enhay), *Hankul* 127.

Lee & Ramsey 2000. *The Korean Language*. Albany:State University of New York Press.

Lee, S-O. 1978. Middle Korean Tonology. PhD dissertation, University of Illinois.

Lehiste, I. 1970. *Suprasegmentals*. Cambridge. MA: MIT Press.

Lehiste, I. and G. Peterson. 1961. Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America* 33, 419-25.

Liberman, M. 1975. The intonational system of English. PhD dissertation, MIT.

Liberman, M., and J. Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In M. Aronoff and R. Oehrle (ed.), *Language Sound Structure*, pp.157-233. Cambridge, MA: MIT Press.

Lin, H.B., and B.Repp. 1989. Cues to the perception of Taiwanese tones. *Language and Speech* 32.1: 25-44.

Lindblom, B., Lyberg, B., and K. Holmgren.1981. *Durational patterns of Swedish phonology: Do they reflect short-term memory processes?* Bloomington: Indiana University Linguistics Club.

Maddieson, I. 1985. Phonetic cues to syllabification, in Victoria Fromkin, ed. *Phonetic Linguistics*. Orlando: Academic Press.

Mohr. B. 1971. Intrinsic variations in the speech signal. *Phonetica* 23, 65-93.

Moon, H-K. 1962. Taegu pangen uy koce cangtan (Pitch and Length in Taegu dialect). *Inmun kwahak*, 7: 53-103.

Moore, C. B., and A. Jongman. 1997. Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America*, 102: 1864-1877.

Myers, S. 1999. Tone Association and F0 timing in Chichewa. *Studies in African Linguistics* 28 (2): 215-239.

2003. F0 timing in Kinyarwanda. *Phonetica* 60: 71-97.

O'Connor, J.D. and G. F. Arnold. 1973. *Intonation of colloquial English.* London: Longman.

Ohala, J and W. Ewan. 1973. Speech of pitch change. *Journal of the Acoustical Society of America 53: 345.*

Oller, D. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54: 1235-1247.

Park, K-C. 2003. The structure of accentual phrase in Korean: the interaction between segments and suprasegments in three Korean dialect. PhD dissertation, Indiana University.

Peterson, G.E. and H.L. Barney. 1952 Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24:75-184.

Peterson, G.E. and I. Lehiste. 1960. Duration of syllable nuclei in English, *Journal of the Acoustical Society of America* 32: 693-703.

Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. PhD dissertation, MIT.

Pierrehumbert, J., and M. Beckman. 1986. Intonational Structure in Japanese and English. *Phonology Yearbook* 3:255-309.

Pierrehumbert, J., and S. Steele. 1989. Categories of Tonal Alignment in English. *Phonetica* 46, 181-196.

Pike, K. 1948. *Tone languages*. Ann Arbor, Michigan: University of Michigan Press.

Prieto, P. et al. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23: 429-251.

2005. Stability effects in tonal clash context in Catalan. *Journal of Phonetics* 33:215-242.

Pulleyblank, D. 1986. *Tone in lexical phonology*. Dordrecht:D. Reidel Publishing Co.

Purcell, E.T. 1976. Pitch peak location and the perception of Serbo-Croatian word tone. *Journal of Phonetics* 4:265-270.

Ramsey, S. R. 1975. Accent and Morphology in Korean Dialect: A descriptive and historical study. PhD dissertation, Yale University.

Rosen, S. 1977. Fundamental frequency patterns and the long-short vowel distinction in Swedish. Speech Transmission Laboratory -Quarterly Progress and Status Report 1:31-37.

Silverman, K., and J. Pierrehumbert. 1990. The timing of prenuclear high accents in English, in Kingston, Beckman. *Papers in Laboratory Phonology I: Between the grammar and the physics of speech*: 72-106.

Sohn, H-M. 1999. *The Korean Language.* Cambridge: Cambridge University Press.

Sproat, R. and O. Fujimura. 1993. Allophonic Variation in English /l/ and its Implications for Phonetic Implementation. *Journal of Phonetics* 21: 291-311.

Steel, S. 1986. Nuclear accent F0 peak location: Effect of rate, vowel, and number of following syllables. *Journal of acoustical society of America* 80. suppl. 1. pp. S51.

Sundberg, J. 1979. Maximum speed of pitch changes in singers and untrained singers. *Journal of Phonetics* 7: 71-79

Sohn, Hyang-Sook, 1987. Underspecification in Korean Phonology. PhD dissertation, University of Illinois at Urbana-Champaign.

This document does not include the vita page from the original.

Um, Y-N. 2000. Tone and Related phonological aspects of Middle Korean and Modern Korean dialects with special reference to Vowel Length. PhD dissertation, University of New York at Buffalo.

Wong, Y. W. 2006. Contextual Tonal Variations and Pitch Targets in Cantonese. In *Proceedings of Speech Prosody 2006*, Dresden, Germany**:** PS3-13-199.

Xu, Y. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25, 61-83.

1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27:55-105.

2001. Fundamental frequency peak delay in Mandarin. *Phonetica* 58: 26-52.

Xu, Y. and X Sun. 2002. Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America* 111: 1399-1413

Yuan, J. 2003. Selective adaptation tests on Mandarin Tone2 and Tone3. *15th International Congress of the Phonetic Sciences*, UAB, Barcelona.

Zsiga, E. 1997. Features, Gestures, and Igbo Vowels: An Approach to the Phonology-Phonetics Interface. *Language* 73 (2): 227-274